PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

Assistant Commissioner for Patents
United States Patent and Trademark Office
Box PCT
Washington, D.C.20231
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 20 March 2000 (20.03.00)	
International application No. PCT/US99/16831	**Applicant's or agent's file reference** 66043
International filing date (day/month/year) 23 July 1999 (23.07.99)	**Priority date** (day/month/year) 24 July 1998 (24.07.98)
Applicant CHUAQUI, Rodrigo, F. et al	

1. The designated Office is hereby notified of its election made:

[X] in the demand filed with the International Preliminary Examining Authority on:

03 February 2000 (03.02.00)

[ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was

[ ] was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer
Nestor Santesso

Telephone No.: (41-22) 338.83.38

PATENT COOPERATION TREATY

REC'D 19 OCT 2000

WIPO PCT

PCT

15

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
66043		
International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US99/16831	23/07/1999	24/07/1998
International Patent Classification (IPC) or national classification and IPC C12N15/12		
Applicant THE GOVERNMENT OF THE UNITED STATES... et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 5 sheets, including this cover sheet.

   ☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☐ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   - VI ☐ Certain documents cited
   - VII ☐ Certain defects in the international application
   - VIII ☐ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
03/02/2000	13.10.2000
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Pilat, D Telephone No. +49 89 2399 8668

Form PCT/IPEA/409 (cover sheet) (January 1994)

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/US99/16831

I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

Description, pages:

1-24 as originally filed

Claims, No.:

1-16 as originally filed

Drawings, sheets:

1/7-7/7 as originally filed

2. The amendments have resulted in the cancellation of:

☐ the description, pages:
☐ the claims, Nos.:
☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	
	No:	Claims	1-7,9-12
Inventive step (IS)	Yes:	Claims	
	No:	Claims	8, 13-16
Industrial applicability (IA)	Yes:	Claims	1-16
	No:	Claims	

2. Citations and explanations

see separate sheet

Ad Section I: Basis of the report

1. Reference is made to the following documents:

D1: COLE ET AL.: 'cDNA sequencing and analysis of POV1 (PB39): a novel gene up-regulated in prostate cancer' GENOMICS, vol. 51, no. 2, July 1998 (1998-07), pages 282-287, XP002120985

D2: WO 98 21328 A (KATO SEISHI ;PROTEGENE INC (JP); SEKINE SHINGO (JP); SAGAMI CHEM R) 22 May 1998 (1998-05-22)

D3: WO 98 10098 A (FISHER PAUL B ;UNIV COLUMBIA (US)) 12 March 1998 (1998-03-12)

D4: CHUAQUI ET AL.: 'Identification of a novel transcript up-regulated in a clinically aggressive prostate carcinoma' UROLOGY, vol. 50, no. 2, August 1997 (1997-08), pages 302-307, XP002120986

D5: HILLIER ET AL.: 'The WashU-Merck EST Project' EMBL ACC NO: R00504, 17 April 1995 (1995-04-17), XP002120987

Ad Section V :Reasoned statement under Rule 66.2(a)(ii); citations and explanations supporting such statement

2. Novelty (Article 33 (2) PCT)

2.1 D1 discloses the complete sequence of POV1 (PB39) which includes a sequence which is identical to the EST R00504. D1 describes that this novel gene was up-regulated in human prostate cancer and contains an open reading frame of 559 amino acids (see abstract). Two transcript were identified which encode a 560 a.a. protein wherein the 47 C terminal amino acid sequence found in one transcript was replaced by 48 new amino acid sequence (see p.282, col.2 last full sentence, Fig.2A and 2B).
Thus, the subject-matter of claims 1 to 7 and 9 to 12 lack novelty.

2.2 The applicant is requested to note that the same objection of lack of novelty may be raised based on the content of D2 to D4 for one or more claims recited above.

D2 document relates to the identification of a transcript that is up-regulated in

aggressive prostate carcinoma. It refers to a 103 bp fragment that has been identified to be up-regulated in tumor cells (see D2 col.1 lines 10-17) which allowed the identification of the EST R00504 sequence (see D3). Finally, D2 describes a method for assessing prostate cancer using this marker (see table I). Finally, D4 discloses p.166 SEQ ID N°71. This sequence shares >99% identity in 2260 bp overlap with SEQ ID N°1.

3. Inventive step (Article 33 (3) PCT)

Claims to antibodies and to a method for detecting precancerous cell or cells by means of detecting up-regulation of a protein, in particular by means of using an antibody are well known in the art (see also D5, claim 13). The combination of this application with either a known amino acid sequence or an obvious antibody (see point 2.1 above) as claimed in claims 8 and 13-16 would have been considered by the skilled person in accordance with circumstances. As a consequence, these claims lack an inventive step.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/47, 16/18, C12Q 1/68, G01N 33/53

A1

(11) International Publication Number: WO 00/05376

(43) International Publication Date: 3 February 2000 (03.02.00)

(21) International Application Number: PCT/US99/16831

(22) International Filing Date: 23 July 1999 (23.07.99)

(30) Priority Data:
60/094,137 24 July 1998 (24.07.98) US

(71) Applicant *(for all designated States except US)*: THE GOVERNMENT OF THE UNITED STATES OF AMERICA, represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; c/o National Institutes of Health, Office of Technology Transfer, Box OTT, Bethesda, MD 20892 (US).

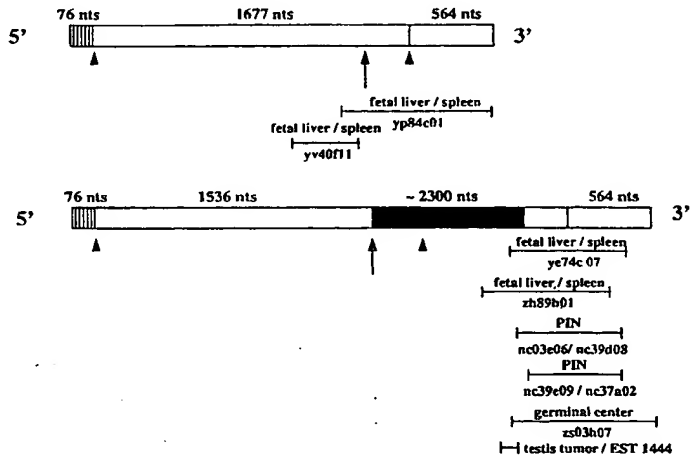
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: CHUAQUI, Rodrigo, F. [CL/US]; Apartment 203, 10630 Montrose Avenue, Bethesda, MD 20814 (US). COLE, Kristina, A. [US/US]; 6011 Wilmett Road, Bethesda, MD 20817 (US). LIOTTA, Lance, A. [US/US]; 8601 Bradley Boulevard, Bethesda, MD 20817 (US).

(74) Agents: SAMPLES, Kenneth, H. et al.; Fitch, Even, Tabin & Flannery, Suite 1600, 120 South LaSalle Street, Chicago, IL 60603–3406 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: PB 39, A GENE DYSREGULATED IN PROSTATE CANCER, AND USES THEREOF



(57) Abstract

A novel gene, PB39, that is up–regulated, or over–expressed, in prostate cancer has been identified. The gene has been identified by means of its cDNA obtained by reverse transcription of the corresponding mRNA. Microdissection of prostate glands that had been surgically removed from prostate cancer patients revealed a novel up–regulated transcript in an aggressive prostate carcinoma. Differential analysis for the presence of this gene was carried out from the same glands by comparing tanscription in microdissected normal prostatic epithelium versus that in microdissected invasive tumor. The transcript was over–expressed in 5 of 10 prostate carcinomas examined. A variant transcript was over–expressed in 4 of 4 prostate carcinomas, and was found in 1 of 4 normal samples. The invention provides a purified and isolated nucleic acid that includes the sequence of PB39 or its complement, the sequence of a variant of PB39 or its complement, and a primer or probe, that includes a sequence that is a fragment of these sequences. Additionally, the polypeptide encoded by these genes, an antibody to the polypeptide, and methods of detection of PB39 or its gene product are provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

PB39, A GENE DYSREGULATED IN PROSTATE CANCER, AND USES THEREOF

FIELD OF THE INVENTION

This invention relates to a novel gene, PB39, and variants thereof, which is dysregulated in prostate cancer. This invention also relates to polypeptides encoded by these genes, antibodies to the polypeptides, and to methods for detection of PB39, variants thereof, and/or its gene product. These methods can be used to assess the presence of precancerous or cancerous cells in the prostate gland.

BACKGROUND OF THE INVENTION

Prostate cancer is the most common malignancy in men and the second leading cause of cancer mortality in the United States (Wingo P. et al., Cancer statistics, Cancer J Clin 45: 8-30 (1995)). The disease can progress with markedly different clinical outcomes. An understanding of the difference between aggressive and nonaggressive tumors at the molecular level has been hindered by the diverse cell types present in the prostate gland and an inability to derive pure cell populations for genetic study.

The Gleason score is a reliable prognosticator for disease progression at both ends of the histologic spectrum (≤ 4 or ≥ 8). Poor disease outcome in the more prevalent Gleason range (5 to 7) is best predicted by a positive surgical margin, capsular penetration, and seminal vesicle invasion (Epstein J. et al., Prediction of progression following radical prostatectomy: a multivariate analysis of 721 men with long-term follow-up, Am J Surg Pathol 20: 286-292 (1996)). However, these parameters can only be assessed after major surgery. Certain cases of prostate cancer respond to surgical and/or pharmaceutical intervention. Others metastasize rapidly, by mechanisms that remain poorly understood, to tissues such as bone.

Currently available diagnostic methods directed to the detection of prostate cancer focus largely on prostate specific antigen (PSA). This protein is known to be a glycosylated serine protease, and is produced in relatively large quantities in the epithelial cells of the

prostate and is secreted in seminal fluid. Smaller amounts are detected in the serum of healthy individuals; higher serum concentrations are thought to be correlated with pathological conditions such as prostate cancer. Other diagnostic methods address nucleic acids differentially expressed in prostate cancer.

U. S. Patent 5,674,682 issued Oct. 7, 1997, to Croce et al. relates to methods for detecting prostate cancer micrometastasis, in which a sample containing nucleic acids is amplified and probed by hybridization to particular oligonucleotide probes. These probes are disclosed as being specific for prostate cancer.

U. S. Patent 5,658,730 issued Aug. 19, 1997, to McGill et al., entitled "Methods of Human Prostate Cancer Diagnosis", discloses diagnostic techniques for the detection of human prostate cancer. A set of degenerate probes is used to detect gene amplification in prostate cancer cells at regions of chromosome 8q24.1-24.2. A comparison of the probe sequence with the specific primers used reveals no significant segments sharing identity between either of the specific primers and the disclosed probe. Copy number changes of chromosome 8q serve as a marker for the development of aggressive prostate cancers.

U. S. Patent 5,622,829 issued Apr. 22, 1997, to King et al., entitled "Genetic Markers for Breast, Ovarian and Prostatic Cancer", discloses the nucleotide sequences for several alleles of *BRCA1*. The specification suggests ascertaining men at risk for prostatic cancer in view of female siblings or family members diagnosed for breast cancer. Several alleles of *BRCA1* are disclosed. A method of screening a patient for prostatic cancer susceptibility based on hybridizing with nucleic acids comprising the sequences provided is also claimed.

U. S. Patent 5,614,372 issued March 25, 1997, to Lilja et al. relates to a bioaffinity assay of PSA using monoclonal antibodies in which a measure of PSA is related to the total of the concentration of PSA plus human glandular kallikrein-1 present in a sample of body fluid. The PSA level determined may be either the concentration of free PSA (i.e., uncomplexed PSA) or the concentration of PSA complexed with alpha-1 anti-chymotrypsin. The assay results permit discrimination between benign prostatic hyperplasia and prostate cancer.

U. S. Patent 5,552,277 issued Sep. 3, 1996, to Nelson et al., entitled "Genetic Diagnosis of Prostate Cancer", teaches that prostatic glutathione-S-transferase promoter becomes hypermethylated in most prostatic cancers. Nelson et al. discloses a method of

detecting the hypermethylated promoter in view of its altered susceptibility to a methylation-sensitive restriction nuclease.

U. S. Patent 5,543,296 issued Aug. 6, 1996, to Sobol et al., entitled "Detection of Carcinoma Metastases by Nucleic Acid Amplification", discloses a method for detecting metastasis of a prostate carcinoma which entails treating a sample of non-prostate tissue in such a way as to amplify mRNA for PSA. A method for detecting carcinoma metastases in body tissues and fluids is disclosed only in general, broad terms; prostatic acid phosphate [sic, phosphatase] and PSA are mentioned only at col. 12, l. 23-26. Several primers are disclosed.

U. S. Patent 5,506,106 issued Apr. 9, 1996, to Croce et al., entitled "Methods of Detecting Micrometastasis of Prostate Cancer", discloses a procedure for diagnosing prostate cancer metastasis by seeking mRNA from PSA among the population of nucleated cells in a blood sample, using RT-PCR with PSA-specific primers.

U. S. Patent 5,501,983 issued Mar. 26, 1996 to Lilja et al., entitled "Assay of Free and Complexed Prostate-Specific Antigen", discloses methods of immunoassay for measuring PSA both free and as a proteinase inhibitor complex.

In spite of these advances, there remains a need to develop prognostic tumor markers that can be measured in limited needle biopsies early in the progression of prostate cancer, especially in the case of aggressive prostate cancer. There further remains a need to address more selective and specific assays for various types of prostate cancer. Assays are needed that permit distinguishing prostate cancer from non-neoplastic prostate disease. Furthermore there is a need to develop diagnostic methods that facilitate identifying prostate cancers of differing aggressiveness and metastatic potential. There is likewise a need for the diagnostic probes on which such methods are founded. This invention addresses these needs.

SUMMARY OF THE INVENTION

The invention relates to a novel gene, PB39, that is up-regulated, or over-expressed, in prostate cancer. The gene has been identified by means of its cDNA obtained by reverse transcription of the corresponding mRNA. Microdissection of prostate glands that had been surgically removed from prostate cancer patients revealed a novel up-regulated transcript in an aggressive prostate carcinoma. Differential analysis for the presence of this gene was carried out from the same glands by comparing transcription in microdissected normal

prostatic epithelium versus that in microdissected invasive tumor. The transcript was over-expressed in 5 of 10 prostate carcinomas examined. A variant transcript was over-expressed in 4 of 4 prostate carcinomas, and was found in only 1 of 4 normal samples.

The invention provides a purified and isolated nucleic acid that includes the sequence of PB39 given in SEQ ID NO:1 or its complement. and the sequence of the PB39 variant given in SEQ ID NO:3 or its complement. In particular embodiments, the nucleic acid may be an RNA or a cDNA. Additionally, the invention provides a purified and isolated nucleic acid, such as a primer or probe, that includes a sequence that is a fragment of the PB39 sequence or its complement, or a fragment of the PB39 variant or its complement. Examples of such primers are given in SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:11, and SEQ ID NO:12.

The invention additionally provides polypeptides, whose sequences are given in SEQ ID NOs:2 and 4, that are encoded by nucleic acids that include either the sequence for PB39 or the sequence for the PB39 variant. In one embodiment, this polypeptide is a recombinantly produced polypeptide. The invention furthermore provides an antibody that binds immunospecifically with PB39 or PB39 variants.

A method of detecting precancerous cells or cancer cells in the prostate of a subject is also provided in this invention. The method includes providing a sample of tissue or fluid from the subject and determining whether the sample contains an abnormally high content of a nucleic acid that includes the sequence of PB39 given in SEQ ID NO:1 or its complement, or the sequence of the PB39 variant given in SEQ ID NO:3 or its complement, or a fragment of these sequences. A finding that the sample contains an abnormally high content of the nucleic acid indicates that the subject has precancerous cells or cancer cells in the prostate. In an important embodiment of this method, the determining step includes amplifying the nucleic acid and detecting the amplified nucleic acid.

Additionally the invention provides a method of detecting precancerous or cancer cells in the prostate of a subject. This method includes providing a sample of tissue or fluid from the subject and determining whether the sample contains an abnormally high content of a polypeptide that is the gene product of the PB39 gene or the variant PB39 gene. Finding that the sample contains an abnormally high content of the polypeptide indicates that the subject has precancerous or cancer cells in the prostate. In an important embodiment of this method, the determining step further includes contacting at least a portion of the sample with

an antibody that binds immunospecifically with the polypeptide and determining the amount of the antibody that has bound with the polypeptide present in the sample. In particular embodiments of both of these methods, the sample may be a body fluid, or may be tissue originating from the prostate.

BRIEF DESCRIPTION OF THE DRAWING

Figure 1. Panel a. Denaturing electrophoresis gel comparing reverse transcription-polymerase chain reaction (RT-PCR) amplification products produced by arbitrary and zinc finger primers. Five tumor samples (T) and two normal samples (N) from NCI Patient 1542 are shown. Two bands increased in the tumor samples are present (arrow and arrowhead). R00504 is represented by the arrow. Panel b. Denaturing electrophoresis gels showing R00504 overexpression in two cases (Table 1, cases 1 and 2). Left: RT-PCR amplification of R00504 in normal (N) and tumor (T) samples. Right: Amplification of beta-actin from the same samples. +, reverse transcriptase reaction with Moloney Murine Leukemia Virus (MMLV); -, MMLV reverse transcriptase was replaced by water; C, positive control.

Figure 2. Northern blot of fetal tissue samples using the specific probe R00504. Lanes contain RNA as follows: lane 1, kidney; lane 2, liver; lane 3, lung; and lane 4, brain. Standards are shown on the right.

Figure 3. Panel A. Clontech adult tissue Northern blots probed with radiolabelled R00504 insert. The exposure times of the blots on the left and right were 40 hrs and 6 hrs, respectively. Panel B. Clontech fetal tissue Northern blot probed with radiolabelled R00504 insert. The exposure time was 40 hrs.

Figure 4. Panels A and B. Nucleotide and amino acid sequence of PB39. The nucleotide sequence is numbered on the left and the amino acid sequence numbered on the right. The underlined ATG start is at nucleotide position 77. For the two sequences beginning at nucleotide position 1613, the upper nucleotide and amino acid sequences refer to the 2.3 kb transcript. The lower nucleotide and amino acids sequences refer to the 5 kb transcript. Panel C. Sequence overlap and divergence between 2.3 kb and 5 kb transcripts (upper and lower, respectively). Open reading frame (between arrowheads), 5' untranslated region (UTR) (vertical fine pattern), 3' UTR (area downstream of arrowhead), inserted sequence of the 5kb

transcript (black). position of divergence (arrow). The white area corresponds to the same sequence in both transcripts. Representative EST clones are identified below each transcript.

Figure 5. Clontech fetal tissue Northern blot probed with radiolabelled 5 kb transcript specific probe. The exposure time was 5 days. The amount of RNA loaded on all of the gels was adjusted to give similar beta-actin hybridization signals and represents approximately 2 μg of polyA-selected mRNA.

Figure 6. Regional localization of PB39. Panel A. Localization to chromosome 11p. 4'6'-Diamino 2-phenyl-indole (DAPI) counterstained metaphase showing the location of the PB39 gene (red, at arrows) on the short arm of chromosome 11. Panel B. "G-banded" chromosomal analysis. An example of the DAPI image of Panel A showing chromosome 11 which was converted to a black and white "G-band" image to show the position of PB39 on the short arm of chromosome 11 close to the centromere. The position of PB39 is identified at 11p11.1-11.2. Panel C. DAPI and simulated G-banded image of chromosome 11 after fluorescence in situ hybridization (FISH). Arrows show location of PB39. Panel D. Ideogram of chromosome 11.

Figure 7. Western blots of PB39 probed with polyclonal antibodies raised against peptides 646 (Panel A), 644 (Panel B) and 656 (Panel C).

DETAILED DESCRIPTION OF THE INVENTION

The gene designated PB39 identified in the present invention includes a nucleotide sequence previously found in expressed sequence tag (EST) R00504 (GenBank). The procedure that led to its discovery and identification depends on an analysis of differential transcription of the gene in neoplastic tissue from a prostate gland that is cancerous compared to normal tissue from the same gland. Specifically, microdissection of epithelial tissue from prostate glands that had been surgically removed from prostate cancer patients was carried out. PB39 in microdissected invasive tumor compared to that in microdissected normal prostatic epithelium was found to be over-expressed in 5 of 10 prostate carcinomas examined. These findings identify the condition detected as prostatic intraepithelial neoplasia, the earliest precursor of prostate cancer. Reliable evaluation of overexpression was aided by comparing transcription to that of the constitutively expressed gene for beta-actin.

The finding that PB39 is over-expressed in prostatic intraepithelial neoplasia indicates that the gene and its gene product are useful in the early diagnosis of this disease, and that they may serve as a marker for its early appearance. Methods that may be employed for this purpose include, without limiting the scope of the invention, assay for the gene by the polymerase chain reaction and in situ hybridization, and analysis for the protein product by immunoassay, or by immunohistochemical detection such as indirect immunofluorescence. Probes and antibody reagents may be developed to permit imaging prostate cancer, both primary and metastatic. Isolation of the PB39 gene and gene product will contribute to the understanding of prostate cancer development and progression, based on experimental studies using methods such as immunohistochemistry or in situ hybridization. The availability of PB39 will additionally permit the development of methods of treatment of subjects determined to have prostate cancer. Specifically, treatment modalities such as chemotherapy, immunotherapy, or antisense nucleotides, for example, may be developed to target the prostate cancer identified by PB39. The gene may be applied to produce the recombinant gene product in appropriate host cells, especially in mammalian cells. The recombinant gene product then may serve as the immunogen to provide antibodies to be applied in the various methods mentioned above.

Initial identification of PB39 has been made as a result of experiments performed on surgically excised cancerous prostate glands. The glands were frozen in liquid nitrogen as rapidly as possible in order to preserve all metabolically unstable species that may be present as well as possible without breakdown. Particular species of interest in the present invention are actively transcribed mRNA species leading to expression of genes that may be characteristic of the cancerous state. RNA was extracted from microdissected invasive tumor cells and from corresponding normal prostatic epithelium from the same gland, and the gene expression profile was examined in the two sets of cells as they exist in vivo.

Microdissection was carried out on sections of the prostate under low power microscopic magnification with sufficient resolution and care to isolate only the types of tissue desired (Chuaqui, R. et al.,. Identification of a Novel Transcript Up-Regulated in a Clinically Aggressive Prostate Carcinoma, Urology 50: 302-307 (1997), incorporated herein by reference). In addition to the microscopic field visualized, guidance was further obtained from adjacent sections that have been treated and stained to enhance the identification of histological features. The dissection procedure was carried out as rapidly as possible in order

to minimize degradation of RNA species by endogenous processes. The microdissection technique is further characterized in Emmert-Buck et al. (Increased gelatinase A and cathepsin B activity in invasive tumor regions of human colon cancer samples, Am J Pathol 145:1285-1290 (1994)) and Zhuang et al. (A microdissection technique for archival DNA analysis of specific cell populations in lesions <1 mm in size, Am J Pathol 146:620-625 (1995)) both of which are incorporated herein by reference.

Each sample of microdissected tissue is then extracted with a procedure, such as the guanidinium thiocyanate-phenol-chloroform method (Chomczynski et al., Single-step of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction, Anal Biochem 162: 156-159 (1987)), in order to provide an RNA sample that may be assessed for the presence of genes of interest. The samples should contain sufficient numbers of cells to provide RNA for the subsequent manipulations. Commonly it is expected that at least about 5,000 cells, and preferably at least about 10,000 cells will be treated for a given sample. If necessary, further purification of the isolated RNA may be carried out. This may include, without intending to limit the purification methods employed, steps such as digestion with deoxyribonuclease, and further purification steps known to workers of skill in cell biology, molecular biology, and cancer research that permit one to isolate mRNA that is actively involved in translation of genes to yield protein products. Such methods are set forth in general in "Current Protocols in Molecular Biology", Ausubel et al., John Wiley and Sons, New York 1987 (updated quarterly), and "Molecular Cloning: A Laboratory Manual", 2nd Ed., Sambrook, Fritsch and Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989, which are incorporated herein by reference.

The purified mRNA sample is then reverse transcribed to yield the corresponding cDNA species using procedures well known to workers of skill in cell biology, molecular biology, and cancer research (Ausubel et al., Sambrook et al.). The cDNA sample is then further analyzed for species that contain sequence motifs known to bind DNA, such as zinc finger motifs. This is carried out by using PCR with primers designed to amplify the motifs sought. In particular, the present invention discloses that one such primer is an arbitrary primer developed by Stratagene (La Jolla, CA), primer A2 (AATCTAGAGCTCCAGCAG (SEQ ID NO:5)), and a zinc finger-directed primer (Zinc 2, GTCGTCGAATTCCACACAGGAGAAAAGCC (SEQ ID NO:6; Stone et al., Targeted

RNA fingerprinting: the cloning of differentially-expressed cDNA fragments enriched for members of the zinc finger gene family, Nucl Acids Res 22: 2612-2618 (1994)).

This PCR generally provides an amplified DNA sample enriched in DNA molecules encoding gene products having zinc finger motifs. The present inventors discovered unexpectedly that differential display gels of PCR products obtained using this primer pair yielded a particular band having pronounced expression in prostatic cancerous epithelium but not in normal epithelium. Upon partial sequencing of this DNA, it was found that a portion of the sequence matched a sequence appearing in the EST database, GenBank accession no. R00504.

Primers specific for R00504 were designed. These are
GCATGTTACAGGTAGAAAAGCC (SEQ ID NO:7) and
CTGGCGTATCTGAAGAGTCTG (SEQ ID NO:8).
These primers may be employed for specific binding to and amplification of sequences contained within R00504, and as such are useful in probing the gene overexpressed in cancerous epithelium. When these R00504-specific probes were labeled and employed in Northern blot analysis of mRNA species from prostatic epithelium samples, a molecule of approximately 2.6 kb was identified. The gene defined in this mRNA transcript is termed PB39. The complete nucleotide sequence for PB39 has been obtained and is provided in SEQ ID NO:1. The sequence of the protein encoded by PB39 is provided in SEQ ID NO:2. The details of the procedures used to obtain this result are provided in Example 6.

It was subsequently determined that a variant of PB39 results from an alternative RNA splicing mechanism during maturation of the RNA transcript. This variant was analyzed by PCR, including using a probe specific for the sequence inserted into the long form. Sequencing of the resulting amplified DNA showed the variant to have a different sequence, and a termination codon yielding a translated gene product one amino acid residue longer than the 2.6 kb form. The nucleotide sequence of the variant PB39 up to its termination codon is provided in SEQ ID NO:3, and the sequence of the protein encoded by this variant is given in SEQ ID NO:4. The details of these procedures and of the results obtained are provided in Example 7.

The transcripts corresponding to SEQ ID NOs:1 and 3 are found to occur in abnormally high concentrations in samples derived from cells of prostate cancer epithelium when referred to the level found in cells from normal epithelium from the same prostate

gland. This observation forms the basis of a method of detecting precancerous cells or cancer cells in the prostate of a subject. The level of PB39 transcript is determined, for example, by using reverse transcription-PCR, Northern blot analysis, or comparable methods known to workers of skill in cell biology, molecular biology, and cancer research. In these methods, the PCR may be carried with any primer pair specific for PB39, including but not limited to, primers that contain the sequences of SEQ ID NOs:7, 8, 10, 11, and 12. The probes to be applied in Northern blot analysis are to be labeled, and may be based on the primers including the sequences of SEQ ID NOs:7, 8, 10, 11, and 12 mentioned above, as well as on labeled forms of nucleic acids containing the PB39 sequences given by SEQ ID NOs:1 and 3 or their complements. The probes for Northern analysis may also employ fragments of the PB39 sequences given by SEQ ID NOs:1 and 3 or their complements, with the proviso that such fragments are specific for PB39 and are long enough to hybridize effectively with the target sequence in the sample being probed. In each case, the levels of the PB39 transcripts are normalized by taking the ratio of the level of the transcript found to the level of the transcript for a constitutive gene present in the same cells. An example of a gene for a constitutively expressed transcript is that for beta-actin. The present invention also provides diagnostic kits including the above novel nucleic acids, primers and probes for purposes of carrying out this method of detection.

The method of detecting precancerous or cancer cells determines whether the level of PB39 transcript is present at an abnormally high level. An "abnormally high" level, or content, of the nucleic acid transcript, as used herein, relates to a ratio of the level of PB39 transcript to the level of beta-actin transcript that is preferably about two times or more higher in the sample of the cancerous epithelium than that found in a set of samples taken from normal epithelium, as expressed by a mean value found therein.

In performing this method, the sample from the subject may be a biopsy sample drawn from the prostate gland of the subject. In favorable cases, such a biopsy may be obtained in a procedure that minimizes invasiveness and discomfort to the subject, such as a needle biopsy. Alternative samples may be a body fluid from the subject, including but not limited to, blood, urine, and seminal fluid. Generally, sampling methods and choices are well known to workers of skill in the art such as urologists and oncologists.

The gene products encoded by the nucleotide sequences of SEQ ID NOs:1 and 3 are provided in SEQ ID NOs:2 and 4, respectively. The proteins incorporating these amino acid

sequences may be produced as recombinant proteins in host cells modified by vectors containing nucleic acid sequences, such as the sequences of SEQ ID NOs:1 and 3, that encode the proteins. Such recombinant proteins may be produced in prokaryotes such as *Escherichia coli*. Preferably, however, eukaryotic hosts will be employed. For example, host cells from members of many families of Lepidoptera, such as SF-9 cells, may be employed. Such host cells are modified to produce the desired protein by infection with a recombinant baculovirus. *Autographa californica*, wherein the recombinant baculovirus carries the gene for the heterologous protein, commonly under the control of the promoter for the gene for the polyhedrin protein of the virus.

Alternatively, various mammalian cells may be used to produce the PB39 protein product upon being transfected with an appropriate vector harboring the gene including the sequence of SEQ ID NO:1 or SEQ ID NO:3. Methods for preparing the vectors and producing the recombinant proteins are set forth broadly in general terms in Ausubel et al. and Sambrook et al., for example. These procedures are well known to workers of skill in the fields of cell biology and molecular biology. The PB39 protein is to be produced as a recombinant protein in a system such as summarized herein, or an equivalent system, and purified to a high degree of purity. Procedures that may be employed in the purification include fractional precipitation, chromatography, centrifugation, and the like. Such procedures are well known to workers of skill in protein chemistry and enzymology and are, for example, set forth in Deutscher, M. P. (ed.), Guide to Protein Purification: Methods in Enzymology, Vol. 182, Academic Press, San Diego, CA, 1990; and Scopes, R. K., Protein Purification: Principles and Practice, 3rd Ed., Springer-Verlag, New York, NY, 1993.

The purified PB39 protein corresponding to SEQ ID NO:2 or SEQ ID NO:4 may be used as the immunogen to prepare antibodies against the respective proteins. The antibody may be a polyclonal antibody, obtained upon immunizing a host such as a rabbit with the PB39 immunogen. Alternatively, one or more monoclonal antibodies may be obtained upon immunizing a host such as a mouse with the PB39 immunogen, and preparing hybridomas that secrete the antibody. Successful hybridomas are obtained as a result of probing hybridoma clones for antibody molecules that bind immunospecifically with the PB39 immunogen. These procedures are well known to workers of skill in the field of molecular immunology, and are set forth in general terms in Harlow and Lane, "Antibodies: A Laboratory Manual", Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1988), and

Coligan. J., A. Kruisbeek, D. Margulies. E. Sevach. and W. Strober, "Current protocols in immunology", John Wiley & Sons. New York (1994), which are incorporated herein by reference.

Antibodies produced according to the procedures described above are useful in immunoassays directed to detecting precancerous cells or cancer cells in the prostate of a subject. In such immunoassays, for example in an enzyme linked immunosorbent assay (ELISA), the concentration of PB39 antigen is detected in a sample obtained from the subject. If the sample is a fluid sample, it may be used as is, or treated to remove cellular components, and furthermore may be volumetrically diluted if necessary to attain an appropriate concentration of the antigen. If the sample is a sample of cells from the prostate gland the sample may be homogenized and the fraction containing the PB39 antigen may be concentrated. If desired, the prostate sample may be microdissected first to provide prostate epithelium. In an ELISA, by way of nonlimiting example, a first antibody that binds immunospecifically with PB39 is immobilized on the surface of a suitable assay vessel, and the remaining surface sites are then blocked with an innocuous protein. The sample suspected of containing PB39 antigen is added and allowed to react with the immobilized first antibody. After thorough rinsing, any immobilized antigen is further treated with a second antibody that binds immunospecifically with PB39, and the second antibody is detected. These procedures are well known to persons of skill in the fields of molecular immunology and diagnostic immunochemistry.

The method of detecting precancerous cells or cancer cells determines whether the level of PB39 antigen is present at an abnormally high level. An "abnormally high" level, or content, of the antigen, as used herein, relates to the level of PB39 antigen that is, preferably, about two times or more higher than that found in a set of samples taken from normal subjects, as expressed by a mean value found therein. These levels or contents may vary, depending on the origin of the particular sample taken. In one embodiment, the level or content is determined as an absolute number representing the concentration or amount of the antigen present in the sample taken. In another embodiment, the level or content of PB39 antigen may be related by ratio to the level or content of a second antigen known to be constitutively present in the particular sample used for the assay. In this embodiment, the value of the ratio of the level of PB39 antigen to the level of the constitutive antigen is considered to be "abnormally high" when it is preferably about two times or more higher than

the ratio found in a set of samples taken from normal subjects or from normal cells, as expressed by a mean value found therein. The present invention also provides diagnostic kits including the above novel antibodies for purposes of carrying out this method of detection.

The predicted N-terminal sequence of the PB39 protein, as shown in SEQ ID NOs:2 and 4, suggests the presence of a signal-recognition particle sequence for a secreted protein. If the PB39 protein is in fact secreted, its concentration may be increased in the serum early in the progression of prostate cancer, in prostatic intraepithelial neoplasia (PIN) for example. Identification of a serum protein characteristic of prostate cancer would be a useful tool for the early detection of this disease. Epidemiologic studies have shown that PIN precedes the development of prostate cancer by several decades in most men. Thus a marker of early malignancy could identify those men who develop PIN lesions early in life and are at greatest risk for developing clinically significant disease (Bostwick, D.G. et al., Molecular biology of prostatic intraepithelial neoplasia, Prostate, 29: 117-134 (1996)).

The PB39 gene described in the present invention is in general dysregulated in disease states such as prostate cancer. It may be subject to altered expression (i.e., overexpression or underexpression) and/or altered processing, resulting in a change in the level of the expressed protein in such a disease state.

EXAMPLES

Example 1. Microdissection of cancerous prostate glands and isolation of RNA.

Patient samples. All tissue samples were obtained from radical prostatectomy specimens from either the Mayo Clinic (Rochester, MN) or the National Cancer Institute (NCI) (Bethesda, MD). Samples were snap-frozen within minutes after surgery and stored at -70°C until use. Unstained 12-μm frozen tissue sections were dissected under microscopic visualization as previously described (Emmert-Buck M. et al., Increased gelatinase A and cathepsin B activity in invasive tumor regions of human colon cancer samples, Am J Pathol 145: 1285-1290 (1994); Zhuang Z. et al., A microdissection technique for archival DNA analysis of specific cell populations in lesions <1 mm in size, Am J Pathol 146: 620-625 (1995)). Essentially pure populations of normal epithelium and invasive tumor were dissected in each case. An adjacent hematoxylin and eosin-stained section was used as a guide to

ensure accuracy of dissection. All dissections were completed within 30 minutes of preparation of the frozen tissue sections.

RNA Isolation. Approximately 5000 to 10,000 normal epithelial or tumor cells were microdissected for each sample. A scaled down version of the Stratagene (La Jolla, CA) RNA Microisolation procedure was used to isolate RNA (Chomczynski P. et al., Single-step RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction, Anal Biochem 162: 156-159 (1987)). After resuspension of the RNA pellet, a DNAse step and re-extraction was performed using the MessageClean™ kit from GenHunter (Nashville, TN) according to the manufacturer's instructions.

Example 2. Reverse Transcription and Polymerase Chain Reaction Amplification of the Prostate-Specific Transcript.

Complementary DNA (cDNA) was obtained by reverse transcription (RT) using the RNAamp™ kit from GenHunter except that 2.5 μM random hexamer primers from Perkin-Elmer (Norwalk, CT) were used instead of the primers supplied. The final mixture was treated as follows: 65°C for 5 minutes, 25°C for 10 minutes, then 1 μL of Moloney Murine Leukemia Virus with reverse transcriptase activity (MMLV) (GenHunter) was added and incubated at 25°C for 10 minutes, 37°C for 40 minutes, and 94°C for 5 minutes. Each RT reaction generated about 20 μL solution containing about 0.5 -1 ng of cDNA. For each RNA sample, a negative control was done for the RT reaction, replacing the MMLV with 1 μL of water.

Arbitrary Zinc Finger Polymerase Chain Reaction (PCR). Several PCR reactions utilizing arbitrary primers from the RAP-PCR kit from Stratagene and degenerate zinc finger primers (Stone B. et al., Targeted RNA fingerprinting: the cloning of differentially-expressed cDNA fragments enriched for members of the zinc finger gene family, Nucl Acid Res 22: 2612-2618 (1994)) were run to assess differences in gene expression between normal epithelium and invasive tumor in NCI patient 1542 (Table 1, case 2; see Example 3). PCR conditions were systematically varied to maximize the reproducibility of bands present on denaturing electrophoresis gels. The specific PCR primers that generated expressed sequence tag (EST) clone R00504 (see below) were: Primer 1: Stratagene arbitrary primer A2,

AATCTAGAGCTCCAGCAG (SEQ ID NO:5),

and Primer 2: zinc 2,

GTCGTCGAATTCCACACAGGAGAAAAGCC (SEQ ID NO:6).

PCR conditions were: 1 cycle of 94°C for 2 minutes, followed by 35 cycles of 94°C for 30 seconds, 50°C for 30 seconds, 72°C for 1 minute, and 1 cycle of 72°C for 10 minutes.

Gel Electrophoresis. Labeled amplified DNA was mixed with an equal volume of formamide loading dye (95% formamide; 20 mM ethylenediaminetetraacetic acid; 0.05% bromophenol blue, and 0.05% xylene cyanol). The samples were denatured for 5 minutes at 94°C and loaded onto a gel consisting of 6% acrylamide (49:1 acrylamide/bis). Bands in the gels were transferred to 3-mm Whatman paper, the paper was dried, and autoradiography was performed with Kodak X-OMAT film.

Results. Normal prostatic epithelial and tumor RNA samples from NCI patient 1542 (Table 1, case 2; see Example 3) were analyzed by the low-stringency RT-PCR procedure outlined above using arbitrary and zinc finger primers. Several parameters, including the identities of the RT primers and PCR primer sets (sizes and sequences), as well as reaction conditions, were varied in an attempt to elicit differences in gene transcription between normal and tumor cells. In general, comparison of normal and tumor samples produced identical patterns of gene expression; several hundred PCR products were observed that did not vary between normal and tumor cells.

However, PCR with primers A2 (SEQ ID NO:5) and zinc 2 (SEQ ID NO:6) resulted in the presence of a strong product selectively in the tumor sample (Figure 1, Panel A). Separate PCR reactions in which only one of the primers was used did not produce a similarly sized band. The band was extracted from the gel, reamplified, and subjected to partial sequencing. Direct sequencing was performed using the Amplicycle™ sequencing kit from Perkin-Elmer according to the manufacturer's instructions. The following 103-base pair (bp) sequence was obtained:

5' ACAGGAATCC CCAGGAGTGA AGAATAAGCA GGAGGCCCCA GATTCACCTT TAGGGCAAGG AGAGAGAAAC AGAGTCAAGT AGGTAGTCAT CTGCCCTTAA GCC 3' (SEQ ID NO:9).

Analysis showed a match of 102 bp out of the 103 bp to a gene sequence in the expressed sequence tag (EST) database (GenBank accession R00504).

The patient from whom this sample was obtained was a 47-year-old black man who first presented with localized (Stage T2A) prostate cancer. Histopathologic examination of the prostatectomy specimen showed a poorly differentiated adenocarcinoma (Gleason score

8). The patient was clinically free of disease until 1 year postoperatively, when he developed rapidly rising prostate-specific antigen levels and clinical evidence of recurrent disease.

Example 3. Analysis of Differential Expression of R00504.

Messenger RNA levels of R00504 were determined in normal prostate epithelium and corresponding invasive tumor cells from a test panel of 10 prostate carcinoma samples (total of 20 samples. Total RNA was recovered from each sample, treated with DNAse, and R00504 was amplified by RT-PCR. The level of the beta-actin gene from each sample, likewise obtained by RT-PCR, was used as an internal standard to quantitate R00504 levels. For the PCR of the beta-actin gene, 1 μL of the cDNA sample was subjected to PCR with specific primers from Clontech (Palo Alto, CA) according to the manufacturer's instructions. All reactions were run at least twice to ensure reproducibility of results, and a control reaction without reverse transcriptase was run in parallel for all normal and tumor samples.

To analyze differential expression, R00504-specific primers were used:

Primer 1 = GCATGTTACAGGTAGAAAAGCC (SEQ ID NO:7);

Primer 2 = CTGGCGTATCTGAAGAGTCTG (SEQ ID NO:8).

PCR conditions using these primers were as follows: 1 cycle of 94°C for 2 minutes, followed by 35 cycles of 94°C for 30 seconds, 60°C for 1 minute, 72°C for 1 minute, and 1 cycle of 72°C for 10 minutes. Beta-actin and R00504 PCR reactions included dilutions of samples to ensure that reactions were not at saturation conditions. All reaction sets were run a minimum of two times to ensure reproducibility of results, and samples were run in parallel with a negative reverse transcriptase control. Overexpression of R00504 was determined visually and was considered to be present when R00504 was selectively expressed in tumor cells or substantially increased in tumors compared with beta-actin expression (Figure 1).

Normal prostatic epithelium and corresponding invasive tumor were microdissected from a test panel of 10 prostate carcinoma samples. Using the specific R00504 PCR primers given by SEQ ID NOs:7 and 8, samples from NCI patient 1542 showed strong expression of R00504 in the tumor sample but no expression in the normal epithelial sample, consistent with the initial experiment utilizing primers A2 and zinc 2 (SEQ ID NOs:5 and 6; see Example 2).

Table 1 presents the results obtained probing the ten patient samples for R00504. Five of the patients in the test panel showed substantial overexpression of R00504 in the

tumor samples (Table 1, cases 1, 2, 6, 9, 10). Figure 1, Panel B shows examples of two cases demonstrating tumor-specific increases in R00504 levels. Case 1 shows expression of R00504 in both the normal and tumor samples, with a relative increase in expression in the tumor. Case 2 shows selective expression in the tumor sample. Overexpression of R00504 in normal cells relative to the corresponding tumor was not observed in any of the cases. These results indicate that R00504 overexpression occurs frequently in prostate cancer. The clinical parameters of the tumors varied among patients (Table 1), and no correlation between R00504 expression and clinical features of the tumors was found. However, the majority of the samples was from patients who had undergone surgery only a short time prior to these studies, so that minimal follow-up data are available.

Example 4. Northern Blot Analysis of R00504 Transcription.

Northern blots were performed to assess the size of the transcript in tissues. In order to serve as a probe, R00504 was amplified and labeled by PCR with sequence-specific primers using ^{32}P-deoxycytidine triphosphate and the PCR conditions described in Example 3. The labeled PCR product was used to probe samples from various human fetal tissues (obtained from Clontech) according to the manufacturer's recommendations. Two million counts per minute per milliliter of hybridization solution was applied.

The Northern blots containing RNA from fetal kidney, liver, lung, and brain tissues showed a single band of approximately 2.6 kilobases (Figure 2). Expression was highest in the fetal liver sample, consistent with the fact that the initial identification of the R00504 EST used a cDNA library from fetal liver.

The full gene represented by the RNA identified in this Example, and which contains the nucleotide sequence provided in Example 2 as part of EST R00504, is termed PB39 herein.

Example 5. Northern blot analysis for PB39 in adult and fetal human tissues.

Northern blot analysis using labeled R00504 as a probe showed a transcript of approximately 2.6 kb which was expressed in tissues of the adult colon, small intestine, ovary, prostate, spleen, and pancreas (Figure 3, Panel A), fetal kidney, liver, and lung (Figure 3, Panel B), and adult liver and skeletal muscle (results not shown). The level of expression was highest in adult pancreas tissue (Figure 3, Panel A).

Example 6. Nucleotide sequence of PB39.

To determine the complete nucleotide sequence of the PB39 cDNA, specific 5' and 3' PB39 primers were generated and used to isolate PB39 specific PCR products from a human pancreas cDNA library using the Rapid Amplification of cDNA Ends (RACE) method (Marathon-Ready cDNA, Clontech). The 3' and 5' PB39 RACE primer sequences GACCGCATAGACTTCTCAGA (SEQ ID NO:10) and GCATGTTACAGGTAGAAAAGCC (SEQ ID NO:7), respectively, were chosen from EST clone R00504. A 700 bp 3' fragment and a 2 kb 5' fragment were produced, subcloned into a plasmid pCR vector, and cycle sequenced. To validate the sequence, gene-specific PCR products amplified from the pancreas library were directly sequenced, and verified by 10 independent sequencing reactions. Assembly of the entire set of sequences produced a 2317 nucleotide cDNA sequence (SEQ ID NO:1) that includes 76 nucleotides of 5' untranslated sequence, a 1677 nucleotide open reading frame (559 amino acids), and 564 nucleotides of 3' untranslated sequence (Figure 4, Panels A and B). The consensus Kozak sequence GCCGCCATGG placed the translation initiation methionine at nucleotide position 77 (Kozak M., Structural features in eukaryotic mRNAs that modulate the initiation of translation, J Biol Chem 266: 19867-19870 (1991)).

Example 7. Nucleotide sequence of a variant of PB39.

Basic Local Alignment Search Tool (BLAST) analysis (Altschul, S. F. et al., Gapped BLAST and PSI-BLAST: A new generation of protein database search programs, Nucleic Acids Res 25:3389-3402 (1997)) of the PB39 sequence against the human EST database showed multiple PB39 EST clones from diverse tissue types. Interestingly, several PB39 homologous clones showed an identical divergence within the coding sequence at nucleotide 1610 (Figure 4, panel B) with introduction of an additional nucleotide sequence. To analyze this longer, alternative form of PB39 further, a PCR primer specific for the inserted sequence (TCTGCAAAGTGGCTGAGATGAG (SEQ ID NO:11)) was designed and used to amplify cDNA from the pancreas library, together with a PB39-specific 5' primer (CCTGCCTTATCTTTCTGAACTGCACC (SEQ ID NO:12)). The amplified product was isolated and directly sequenced. The sequence is shown in Figure 4, panels A and B (SEQ ID NO:3). Analysis of the open reading frame shows the addition of 48 new amino acids beginning at nucleotide position 1613, followed by a stop codon. Thus, the larger transcript

of PB39 encodes a 560 amino acid protein in which the 47 C-terminal amino acids found in the 2.3 kb PB39 are replaced by 48 new amino acids (Figure 4, panels A and B). A schematic diagram of the relationship between the two RNA species is shown in Figure 4, Panel C.

Northern blot analysis using a probe specific for the inserted sequence showed a 5kb PB39 transcript expressed in adult pancreas and fetal liver tissue. As expected, this probe did not hybridize to the 2.6 kb PB39 (Figure 5). A longer exposure of the R00504-probed blots did reveal a less intense transcript at 5kb, which would be expected since this sequence is common to both transcripts (Figure 3, panels A and B).

RT-PCR analysis was performed to study the expression of the 5 kb PB39 transcript in human prostate tissue as described above (Example 3). RT-PCR using primers directed against the inserted sequence in microdissected normal and invasive prostate epithelium showed a product in 4 of 4 tumor samples, but only 1 of 4 corresponding normal samples (results not shown). One of the cases over-expressing the 5 kb transcript did not show over-expression of the 2.6 kb form of PB39. Since the results

TABLE I. *Clinical and histopathologic aspects of test panel of prostate cancer cases*

Case No./Age (yr)	R00504	Gleason Score*	Tumor Stage*	PSA Level (ng/mL)*
1/66	++	4/4	T2A	2.2
2/47	++	4/4	T2C	14.3
3/55		4/3	T3C	49.8
4/66		2/3	T2C	6.3
5/65		3/3	T2C	11.1
6/68	++	4/4	T3A	7.1
7/68		4/3	T3A	49.2
8/54		4/4	T1C	7.7
9/64	++	4/3	T3A	14.1
10/51	++	3/3	T2B	14.9

Key: ++ = increased levels of R00504 in invasive tumor compared with corresponding normal epithelium.
** At diagnossis.*

in Table 1 show overexpression of PB39 in 5 of 10 tumor samples. it appears that both the 2.6 kb and 5 kb forms of PB39 are up-regulated in unique subsets of tumors. The physiological significance of this finding is not yet clear.

The 5 kb PB39 variant transcript was found to match EST clones primarily from fetal and tumor tissue libraries (Figure 4, panel C). This transcript was also found to be highly expressed in a cDNA library from a microdissected prostatic intraepithelial neoplasia focus which was sequenced as part of the Cancer Genome Anatomy Project. Thus, PB39 represents one of the first identified genes whose expression has been shown to be increased early in prostate cancer development. The cellular regulation of PB39 mRNA splice variants, their precise expression levels during prostate tumorigenesis, and the functional significance of altering the C-terminal 47 amino acids remain to be determined.

Example 8. Localization of PB39 on chromosome 11.

A Whitehead Institute Sequence Tagged Site marker WI-17004 (GenBank acc. G22380) maps PB39 to 291.1 cR from the telomere of the short arm of chromosome 11 (Schuler GD, Sequence mapping by electronic PCR, Genome Res 7: 541-550 (1997)). To confirm this result, in situ hybridization was performed following protocols previously described (Pinkel, D. et al., Fluorescence in situ hybridization with human chromosome-specific libraries: detection of the trisomy 21 and translocations of chromosome 4, Proc Natl Acad Sci USA. 85:9138-9142 (1988); Hirai, M. et al., A method for simultaneous detection of fluorescent G-bands and in situ hybridization signals, Cytogenet Cell Genet 66: 149-151 (1994)). The chromosomal localization of the gene was determined by hybridization of the 2 kb 5' RACE PCR product to metaphase chromosomes and converting DAPI banding to "G-banding" using IP Lab Spectrum™ (Scan Analytics, Fairfax, VA) software for chromosomal identification. This result indicates that the PB39 gene maps to human chromosome 11p11.1-11.2. A total of 50 cells was examined to determine the precise chromosomal location of the probe. In all metaphases scored, clear signals were seen on the short arm of chromosome 11 (Figure 6, panels A and B). DAPI banding unambiguously showed the position of the signal in the region 11p11.1-11.2 (Figure 6, panels A, B, C and D). Both the STS primers and the fluorescence in situ hybridization (FISH) probe were directed against sequences common to the 2.6 kb and 5kb transcripts. In both cases hybridization to only one chromosomal location was identified. Interestingly, the human chromosome 11p11-11p12 region has been

postulated to harbor one or more metastasis suppressor genes, including KAI1, and has also been shown to be deleted in 70% of advanced prostate cancers (Dong, J.T. et al., KAI1, A metastasis suppressor gene for prostate cancer on human chromosome 11p11, Science 268:884-886 (1995); Kawana, Y. et al., Location of KAI1 on the short arm of human chromosome 11 and frequency of allelic loss in advanced prostate cancer, Prostate, 32: 205-213 (1997)).

Example 9. Expression of PB39 in COS cells.

A cDNA encoding PB39 is to be ligated into a vector that can be used to transfect a mammalian cell. In particular, SV40 vectors may be employed for this purpose. The PB39 gene is to be subcloned into an appropriate plasmid such that suitable restriction sites become available. The cDNA is then to be extracted using the appropriate restriction enzyme, and ligated into a complementary site in the SV40 genome.

The recombinant SV40 is to be used to transfect, for example, simian COS cells. The vector is combined with calcium phosphate for incorporation into the cells. Amplification using vectors incorporating the drug resistance gene *dhfr* permits larger amounts of the cloned gene to accumulate upon exposure of the culture to methotrexate, and to be expressed in the cultured cells. Upon completion of the selection, the amplified cells are to be cultured to produce the recombinant PB39 gene product.

PB39 protein is to be isolated and purified from the cultured cells. The cells are to be harvested, and gently disrupted to release the cellular contents into the medium. PB39 is to be purified using known methods including differential centrifugation, and various forms of column chromatography, such as ion exchange, hydrophobic interaction, and size exclusion. Purity is to be assessed using analytical chromatographic procedures, sodium dodecyl sulfate-polyacrylamide gel electrophoresis and/or like procedures.

Example 10. Production of polyclonal anti-PB39 antibodies.

Putative epitopic peptide sequences were synthesized in order to elicit polyclonal antibodies specific for PB39. Peptides common to both the 2.6 kb encoded protein and the variant transcript protein are:

peptide 644: Thr Gln Asp Glu Gln Arg Arg Trp Pro Gly Cys Asp Gln Gln (SEQ ID NO:13),

peptide 645: Glu Asn Leu Pro Glu Arg Ser Val Pro Leu Arg Lys Ser Leu (SEQ ID NO:14), and

peptide 646: Arg Pro Arg Tyr Cys Lys Ile Gln Lys Leu Thr Asn Ala(SEQ ID NO:15).

A peptide sequence specific for the 2.6 kb encoded protein is

peptide 655: Ala Asn Gly Met Gly Pro Leu Lys Val Leu Ser Gly Ser (SEQ ID NO:16).

A peptide sequence specific for the variant transcript protein is

peptide 656: Ala Arg Gly Thr Ser Glu Val Ser Asn Leu Gln Val Ser (SEQ ID NO:17).

Rabbits were immunized with a BSA conjugate of each of the above peptides, using complete Freund's adjuvant. Immunizations were done every two weeks. The rabbits were bled prior to the immunizations and again during the intervals between injections.

The anti-peptide antibody sera were affinity purified over the corresponding peptide-conjugated Affi-Gel 10™ columns. After desorption from the column, the antibody-containing eluates were concentrated for storage and use.

The binding of the resulting antibodies was assessed in western blots. For testing of antibodies raised against peptides 655 and 656 30 μg of protein extracted from a prostate cancer cell line (1542/8.4 clone) were analyzed. For testing of antibodies raised against peptides 644, 645 and 646, protein extracted from roughly 24,000 cells from microdissected tumor samples were analyzed. Proteins were separated on a 8-12% gradient P.A.G.E. gel (NOVEX) run in tris glycine and transferred to a PVDF membrane. The membranes were blocked and incubated overnight at 4°C with varying titers of the anti-peptide antibody mixtures (range from 1:10 to 1:5000, volume/volume). The membranes were washed and incubated with an HRP conjugated goat anti-rabbit secondary antibody. After washing, the membranes were incubated with a chemiluminescence substrate and visualized on X-OMAT film. For all antibodies tested, a 60 kD band was evident. This corresponds to the size anticipated for both the 2.6 kb encoded protein and the variant transcript. In Figure 7, Panels A, B, and C, show the results obtained with the anti-646, -644 and -656 antibodies, respectively. For the westerns probed with anti-655 and -645 antibodies, the bands were present, but were too faint to reproduce in a photograph. Antibody specificity was demonstrated by eliminating the bands upon competition with increasing concentrations of the corresponding epitopic peptide.

Example 11. Monoclonal antibodies.

Purified PB39, or BSA-conjugated peptides 646, 644 or 656, are to be injected into mice together with Freund's complete adjuvant. Subsequently spleen cells obtained from the mice are to be fused with immortalized murine tumor cells to produce hybridomas. Each hybridoma cell is to be expanded, and the supernatant from the culture of each clone is to be assessed for its ability to bind immunospecifically with authentic PB39 antigen. Clones identified as successfully secreting binding antibodies are to be retained and stored. Monoclonal anti-PB39 antibodies from such clones are obtained by culturing these clones.

Example 12. ELISA assay for PB39 in prostate tissue.

Prostate tissue is to be obtained from a subject. If desired, the tissue is to be microdissected to extract cancerous and normal epithelial cells. The tissue sample is to be homogenized. The wells of a microtiter plate are to be contacted with polyclonal anti-PB39 from Example 10 to adsorb the antibody, and remaining sites blocked with casein. Portions of the homogenate are then to be added to the wells of the plate. Aliquots of authentic PB39 antigen at known concentrations are to be added to additional wells to serve as standards for quantitation of the amount of antigen. The wells are then to be treated with either a monoclonal anti-PB39 antibody from Example 10, or with a further aliquot of the polyclonal antibody. This second antibody will previously have been conjugated with horse radish peroxidase. The wells are then treated with the chromogenic substrate *o*-phenylenediamine dihydrochloride, and the resulting color is to be determined on a microtiter plate spectrophotometric detector. The absorbance values so obtained are calibrated against the PB39 standard for quantitating the amount of antigen present in the samples. Of course, other visualizing or detecting systems known in the art can be used if desired.

We claim:

1. A purified and isolated nucleic acid comprising a sequence given in SEQ ID NO:1, a sequence which is complementary to the sequence given in SEQ ID NO:1, a sequence given in SEQ ID NO:3, or a sequence which is complementary to the sequence given in SEQ ID NO:3.

2. The nucleic acid described in claim 1, wherein the nucleic acid is an RNA.

3. The nucleic acid described in claim 1, wherein the nucleic acid is a cDNA.

4. A purified and isolated nucleic acid comprising a sequence that is a fragment of the sequence given in SEQ ID NO:1, the sequence which is complementary to the sequence given in SEQ ID NO:1, the sequence given in SEQ ID NO:3, or the sequence which is complementary to the sequence given in SEQ ID NO:3, wherein the fragment hybridizes specifically with the sequence given in SEQ ID NO:1, the sequence which is complementary to the sequence given in SEQ ID NO:1, the sequence given in SEQ ID NO:3, or the sequence which is complementary to the sequence given in SEQ ID NO:3.

5. The nucleic acid described in claim 4, wherein the nucleic acid is selected from the group consisting of SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:11, and SEQ ID NO:12.

6. A polypeptide encoded by a nucleic acid comprising the sequence given in SEQ ID NO:1 or the sequence given in SEQ ID NO:3.

7. The polypeptide described in claim 6, wherein the polypeptide is a recombinantly produced polypeptide.

8. An antibody that binds immunospecifically with a polypeptide encoded by a nucleic acid comprising a sequence given in SEQ ID NO:1 or a sequence given in SEQ ID NO:3.

9. A method of detecting precancerous cells or cancer cells in the prostate of a subject, said method comprising providing a sample of tissue or fluid from the subject and determining whether the sample contains an abnormally high content of a nucleic acid comprising a sequence given in SEQ ID NO:1, a sequence which is complementary to the sequence given in SEQ ID NO:1, a sequence given in SEQ ID NO:3, a sequence which is complementary to the sequence given in SEQ ID NO:3, or a fragment of any of the sequences, whereby determining that the sample contains an abnormally high content of the nucleic acid indicates that the subject has precancerous cells or cancer cells in the prostate.

10. The method described in claim 9, wherein the sample is a body fluid.

11. The method described in claim 9, wherein the sample is tissue originating from the prostate.

12. The method described in claim 9, wherein the determining step comprises amplifying the nucleic acid and detecting the amplified nucleic acid.

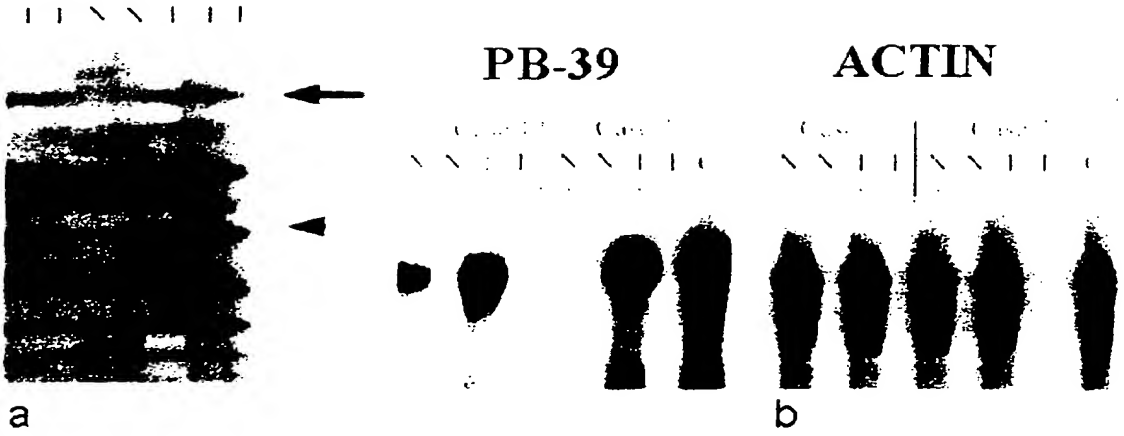
13. A method of detecting precancerous cells or cancer cells in the prostate of a subject, said method comprising providing a sample of tissue or fluid from the subject and determining whether the sample contains an abnormally high content of a polypeptide encoded by a nucleic acid comprising a sequence given in SEQ ID NO:1 or SEQ ID NO:3, whereby determining that the sample contains an abnormally high content of the polypeptide indicates that the subject has precancerous cells or cancer cells in the prostate.

14. The method described in claim 13, wherein the sample is a body fluid.

15. The method described in claim 13, wherein the sample is tissue originating from the prostate.

16. The method described in claim 13, wherein the determining step further comprises contacting at least a portion of the sample with an antibody that binds immunospecifically with the polypeptide and determining the amount of the antibody that has bound with the polypeptide present in the sample.

Figure 1



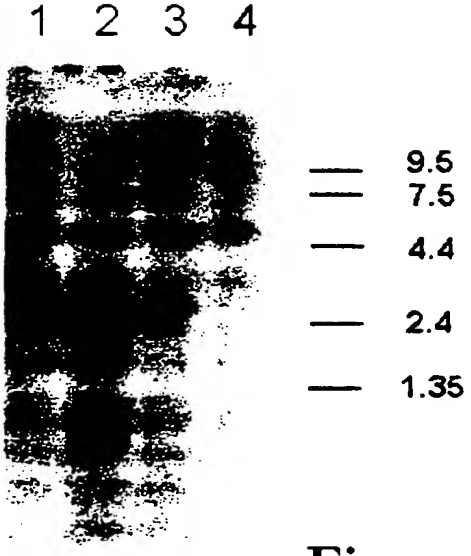


Figure 2

Figure 3A

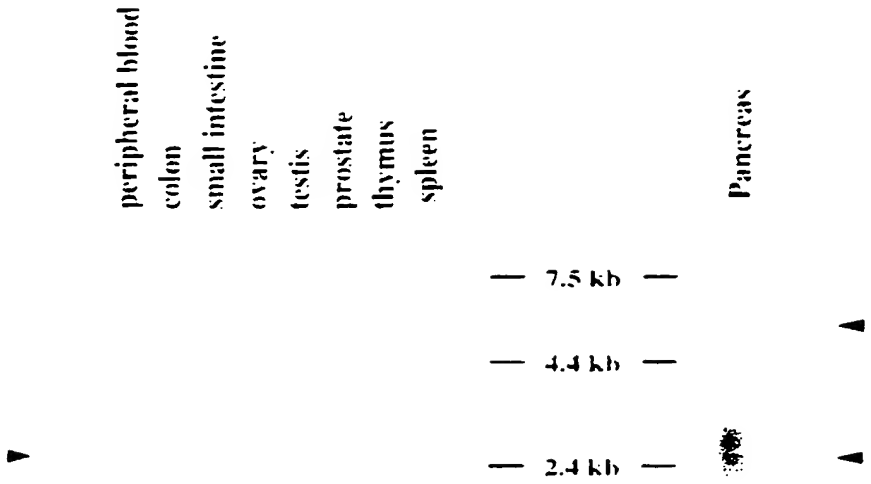


Figure 3B

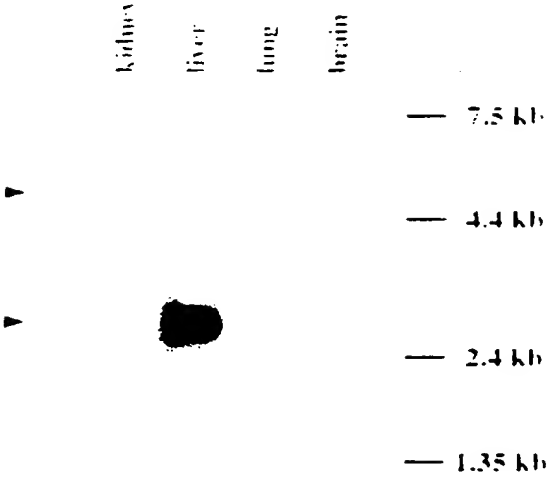


Fig. 4
Panel A


```
   1 ccggggctggaggggggcaagcgggttccgaggtgcaaagcctgg
     tgccccgagccctgcggagctcggggccagc
  77 atggcccccacgctgcaacaggcgtaccggaggcgctggtggatg
      M  A  P  T  L  Q  Q  A  Y  R  R  R  W  W  M   15
 122 gcctgcacggctgtgctggagaacctcttcttctctgctgtactc
      A  C  T  A  V  L  E  N  L  F  F  S  A  V  L   30
 167 ctgggctggggctccctgttgatcattctgaagaacgagggcttc
      L  G  W  G  S  L  L  I  I  L  K  N  E  G  F   45
 212 tattccagcacgtgcccagctgagagcagcaccaacaccacccag
      Y  S  S  T  C  P  A  E  S  S  T  N  T  T  Q   60
 257 gatgagcagcgcaggtggccaggctgtgaccagcaggacgagatg
      D  E  Q  R  R  W  P  G  C  D  Q  Q  D  E  M   75
 302 ctcaacctgggcttcaccattggttccttcgtgctcagcgccacc
      L  N  L  G  F  T  I  G  S  F  V  L  S  A  T   90
 347 accctgccactggggatcctcatggaccgctttggcccccgaccc
      T  L  P  L  G  I  L  M  D  R  F  G  P  R  P  105
 392 gtgcggctggttggcagtgcctgcttcactgcgtcctgcaccctc
      V  R  L  V  G  S  A  C  F  T  A  S  C  T  L  120
 437 atggccctggcctcccgggacgtggaagctctgtctccgttgata
      M  A  L  A  S  R  D  V  E  A  L  S  P  L  I  135
 482 ttcctggcgctgtccctgaatggctttggtggcatctgcctaacg
      F  L  A  L  S  L  N  G  F  G  G  I  C  L  T  150
 527 ttcacttcactcacgctgcccaacatgtttgggaacctgcgctcc
      F  T  S  L  T  L  P  N  M  F  G  N  L  R  S  165
 572 acgttaatggccctcatgattggctcttacgcctcttctgccatt
      T  L  M  A  L  M  I  G  S  Y  A  S  S  A  I  180
 617 acgttcccaggaatcaagctgatctacgatgccggtgtggccttc
      T  F  P  G  I  K  L  I  Y  D  A  G  V  A  F  195
 662 gtggtcatcatgttcacctggtctggcctggcctgccttatcttt
      V  V  I  M  F  T  W  S  G  L  A  C  L  I  F  210
 707 ctgaactgcaccctcaactggcccatcgaagcctttcctgccctt
      L  N  C  T  L  N  W  P  I  E  A  F  P  A  P  225
 752 gaggaagtcaattacacgaagaagatcaagctgagtgggctggcc
      E  E  V  N  Y  T  K  K  I  K  L  S  G  L  A  240
 797 ctggaccacaaggtgacaggtgacctcttctacacccatgtgacc
      L  D  H  K  V  T  G  D  L  F  Y  T  H  V  T  255
 842 accatgggccagaggctcagccagaaggcccccagcctggaggac
      T  M  G  Q  R  L  S  Q  K  A  P  S  L  E  D  270
 887 ggttcggatgccttcatgtcaccccaggatgttcggggcacctca
      G  S  D  A  F  M  S  P  Q  D  V  R  G  T  S  285
 932 gaaaaccttcctgagaggtctgtcccttacgcaagagcctctgc
      E  N  L  P  E  R  S  V  P  L  R  K  S  L  C  300
 977 tcccccactttcctgtggagcctcctcaccatgggcatgacccag
      S  P  T  F  L  W  S  L  L  T  M  G  M  T  Q  315
1022 ctgcggatcatcttctacatggctgctgtgaacaagatgctggag
      L  R  I  I  F  Y  M  A  A  V  N  K  M  L  E  330
1067 taccttgtgactggtggccaggagcatgagacaaatgaacagcaa
      Y  L  V  T  G  G  Q  E  H  E  T  N  E  Q  Q  345
1112 caaaaggtggcagagacagttgggttctactcctccgtcttcggg
      Q  K  V  A  E  T  V  G  F  Y  S  S  V  F  G  360
1157 gccatgcagctgttgtgccttctcacctgccccctcattggctac
      A  M  Q  L  L  C  L  L  T  C  P  L  I  G  Y  375
```

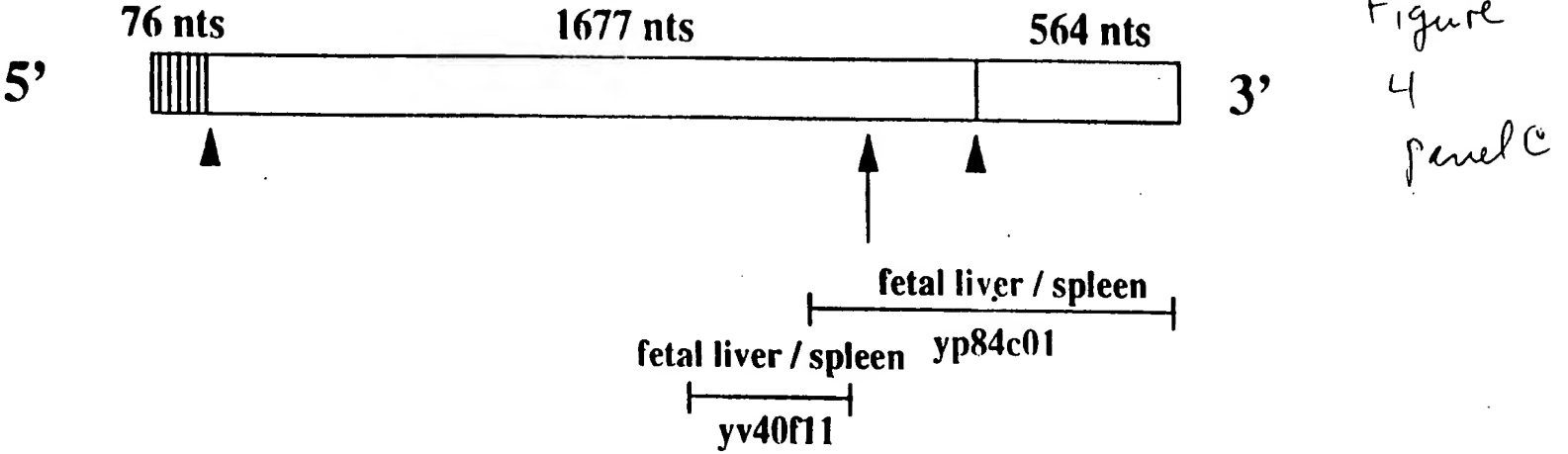
Figure 4 panel B

```
1202 atcatggactggcggatcaaggactgcgtggacgccccaactcag
      I  M  D  W  R  I  K  D  C  V  D  A  P  T  Q  390
1247 ggcactgtcctcggagatgccagggacggggttgctaccaaatcc
      G  T  V  L  G  D  A  R  D  G  V  A  T  K  S  405
1292 atcagaccacgctactgcaagatccaaaagctcaccaatgccatc
      I  R  P  R  Y  C  K  I  Q  K  L  T  N  A  I  420
1337 agtgccttcaccctgaccaacctgctgcttgtgggttttggcatc
      S  A  F  T  L  T  N  L  L  L  V  G  F  G  I  435
1382 acctgtctcatcaacaacttacacctccagtttgtgacctttgtc
      T  C  L  I  N  N  L  H  L  Q  F  V  T  F  V  450
1427 ctgcacaccattgttcgaggtttcttccactcagcctgtgggagt
      L  H  T  I  V  R  G  F  F  H  S  A  C  G  S  465
1472 ctctatgctgcagtgttcccatccaaccactttgggacgctgaca
      L  Y  A  A  V  F  P  S  N  H  F  G  T  L  T  480
1517 ggcctgcagtccctcatcagtgctgtgttcgccttgcttcagcag
      G  L  Q  S  L  I  S  A  V  F  A  L  L  Q  Q  495
1562 ccacttttcatggcgatggtgggacccctgaaaggagagcccttc
      P  L  F  M  A  M  V  G  P  L  K  G  E  P  F  510
1607 tgggtgaatctgggcctcctgctattctcactcctgggattcctg
           agagcgagggttggtgtgggggagcaggagccactctc
      W  V  N  L  G  L  L  L  F  S  L  L  G  F  L  525
            R  A  R  V  G  V  G  G  A  G  A  T  L  525
1652 ttgccttcctacctcttctattaccgtgcccggctccagcaggag
     ctgggggcaggggtagggccttgtatgtggtgccatccctcactc
      L  P  S  Y  L  F  Y  Y  R  A  R  L  Q  Q  E  540
      L  G  A  G  V  G  P  C  M  W  C  H  P  S  L  540
1697 tacgccgccaatgggatgggcccactgaaggtgcttagcggctct
     atctcagccagaggcacctcagaggtctctaatctgcaggtttcc
      Y  A  A  N  G  M  G  P  L  K  V  L  S  G  S  555
      I  S  A  R  G  T  S  E  V  S  N  L  Q  V  S  555
1742 gaggtgaccgcatag     1756
     aagttgtctgccttttag  1759
      E  V  T  A  * 559
      K  L  S  A  F  *#560

     acttctcagaccaagggacctggatgacaggcaatcaaggcctga
     gcaaccaaaaggagtgccccatatggcttttctacctgtaacatg
     cacatagagccatggccgtagatttataaataccaagagaagttc
     tatttttgtaaagactgcaaaaaggaggaaaaaaaaccttcaaaa
     acgccccctaagtcaacgctccattgactgaagacagtccctatc
     ctagaggggttgagctttcttcctccttgggttggaggagaccag
     ggtgcctcttatctccttctagcggtctgcctcctggtacctctt
     gggggggatcggcaaacaggctacccctgaggtcccatgtgccatg
     agtgtgcacaacatgcaatgtgtctgtgtatgtgtgccatgaatg
     tgagaaaaacacagccctcctttcagaaggaaaggggcctgaggg
     ctgtgtcctgggttaggggttgggggtcggccccttccagggcca
     ggaaggcaggttccctctctggtgctgctgcttgcaagtcttaga
     ggaaataaaaagggaagtgag aaaaaaaaa
```

#there are approximately an additional 2300 nucleotides at this site contributing to the 5 kb transcript's 3'UTR.

Figure 4 panel C



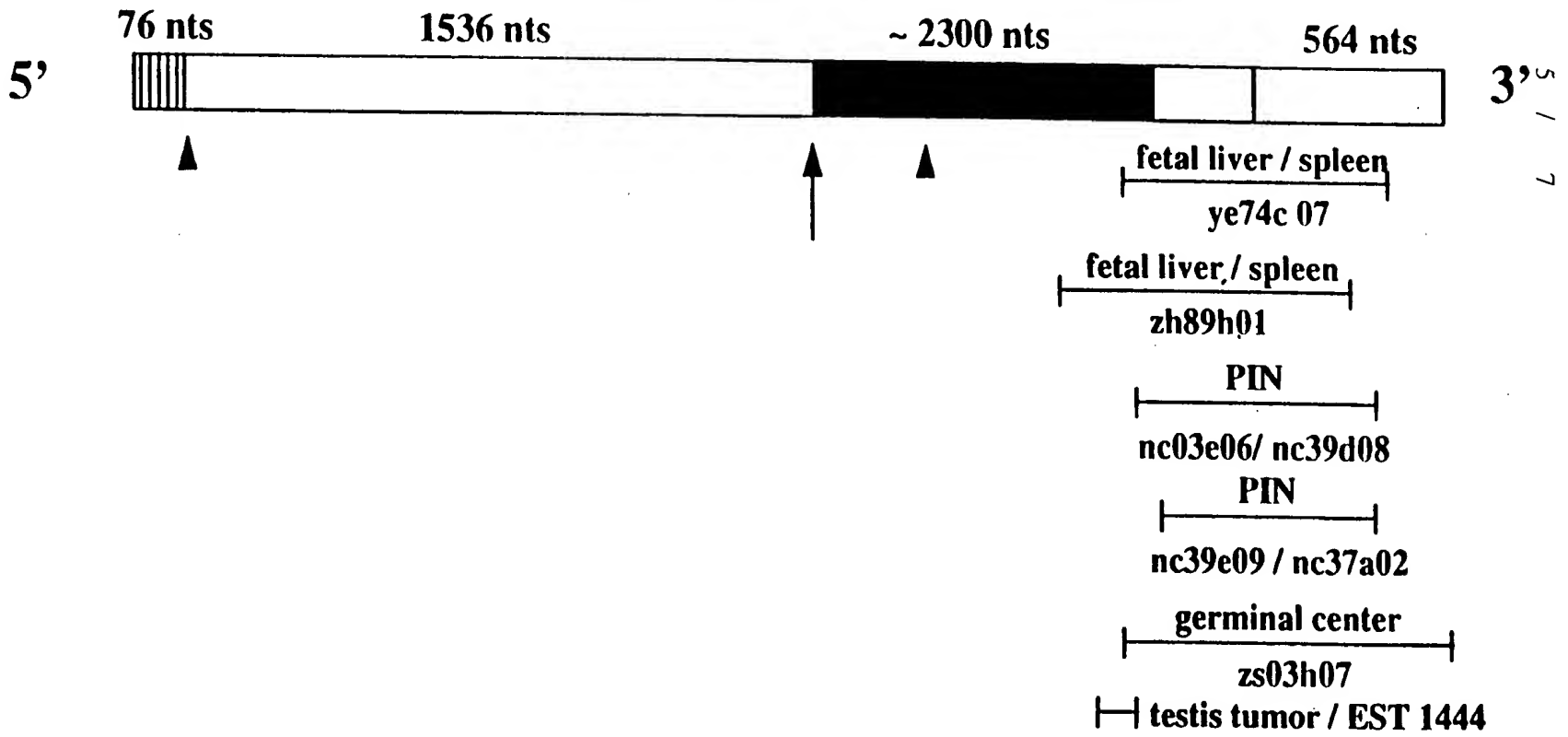


Figure 5

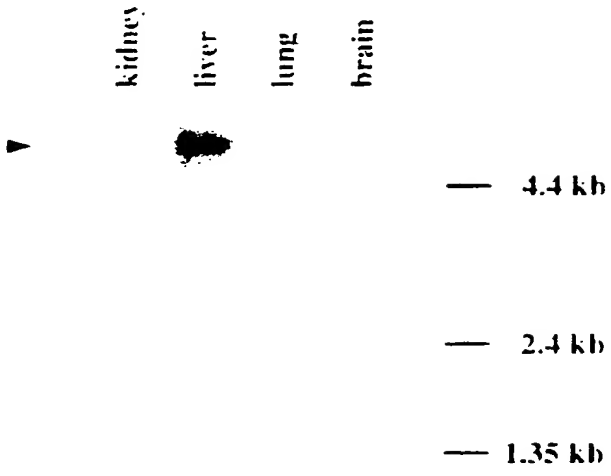


Figure 6

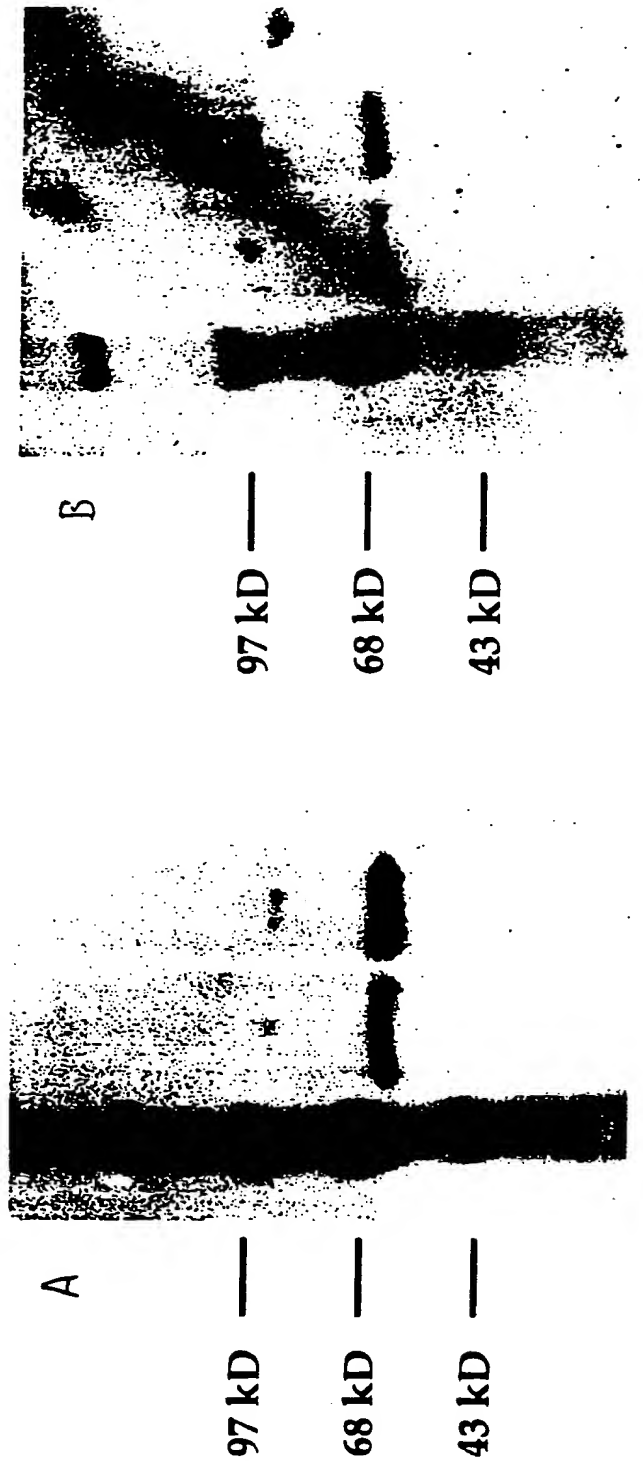

A
97 kD
68 kD
43 kD
Ab 646
B
97 kD
68 kD
43 kD
Ab 644
C
60 kD
42 kD
Ab 656

FIG. 7

SEQUENCE LISTING

<110> Chuaqui, Rodrigo F.
Cole, Kristina A.
Liotta, Lance A.

<120> PB39, A Gene Overexpressed in Prostate Cancer, and Uses Thereof

<130> 6299/62051

<160> 17

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 2326
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (77)...(1753)

<400> 1

```
ccggggctgg agggggggcaa gcgggttccg aggtgcaaag cctggtgccc cgagccctgc      60
ggagctcggg gccagc atg gcc ccc acg ctg caa cag gcg tac cgg agg cgc      112
                  Met Ala Pro Thr Leu Gln Gln Ala Tyr Arg Arg Arg
                   1               5                  10

tgg tgg atg gcc tgc acg gct gtg ctg gag aac ctc ttc ttc tct gct      160
Trp Trp Met Ala Cys Thr Ala Val Leu Glu Asn Leu Phe Phe Ser Ala
         15                  20                  25

gta ctc ctg ggc tgg ggc tcc ctg ttg atc att ctg aag aac gag ggc      208
Val Leu Leu Gly Trp Gly Ser Leu Leu Ile Ile Leu Lys Asn Glu Gly
     30                  35                  40

ttc tat tcc agc acg tgc cca gct gag agc agc acc aac acc acc cag      256
Phe Tyr Ser Ser Thr Cys Pro Ala Glu Ser Ser Thr Asn Thr Thr Gln
 45                  50                  55                  60

gat gag cag cgc agg tgg cca ggc tgt gac cag cag gac gag atg ctc      304
Asp Glu Gln Arg Arg Trp Pro Gly Cys Asp Gln Gln Asp Glu Met Leu
                 65                  70                  75

aac ctg ggc ttc acc att ggt tcc ttc gtg ctc agc gcc acc acc ctg      352
Asn Leu Gly Phe Thr Ile Gly Ser Phe Val Leu Ser Ala Thr Thr Leu
             80                  85                  90

cca ctg ggg atc ctc atg gac cgc ttt ggc ccc cga ccc gtg cgg ctg      400
Pro Leu Gly Ile Leu Met Asp Arg Phe Gly Pro Arg Pro Val Arg Leu
         95                  100                 105

gtt ggc agt gcc tgc ttc act gcg tcc tgc acc ctc atg gcc ctg gcc      448
```

```
Val Gly Ser Ala Cys Phe Thr Ala Ser Cys Thr Leu Met Ala Leu Ala
    110                 115                 120

tcc cgg gac gtg gaa gct ctg tct ccg ttg ata ttc ctg gcg ctg tcc      496
Ser Arg Asp Val Glu Ala Leu Ser Pro Leu Ile Phe Leu Ala Leu Ser
125                 130                 135                 140

ctg aat ggc ttt ggt ggc atc tgc cta acg ttc act tca ctc acg ctg      544
Leu Asn Gly Phe Gly Gly Ile Cys Leu Thr Phe Thr Ser Leu Thr Leu
                145                 150                 155

ccc aac atg ttt ggg aac ctg cgc tcc acg tta atg gcc ctc atg att      592
Pro Asn Met Phe Gly Asn Leu Arg Ser Thr Leu Met Ala Leu Met Ile
            160                 165                 170

ggc tct tac gcc tct tct gcc att acg ttc cca gga atc aag ctg atc      640
Gly Ser Tyr Ala Ser Ser Ala Ile Thr Phe Pro Gly Ile Lys Leu Ile
        175                 180                 185

tac gat gcc ggt gtg gcc ttc gtg gtc atc atg ttc acc tgg tct ggc      688
Tyr Asp Ala Gly Val Ala Phe Val Val Ile Met Phe Thr Trp Ser Gly
    190                 195                 200

ctg gcc tgc ctt atc ttt ctg aac tgc acc ctc aac tgg ccc atc gaa      736
Leu Ala Cys Leu Ile Phe Leu Asn Cys Thr Leu Asn Trp Pro Ile Glu
205                 210                 215                 220

gcc ttt cct gcc cct gag gaa gtc aat tac acg aag aag atc aag ctg      784
Ala Phe Pro Ala Pro Glu Glu Val Asn Tyr Thr Lys Lys Ile Lys Leu
                225                 230                 235

agt ggg ctg gcc ctg gac cac aag gtg aca ggt gac ctc ttc tac acc      832
Ser Gly Leu Ala Leu Asp His Lys Val Thr Gly Asp Leu Phe Tyr Thr
            240                 245                 250

cat gtg acc acc atg ggc cag agg ctc agc cag aag gcc ccc agc ctg      880
His Val Thr Thr Met Gly Gln Arg Leu Ser Gln Lys Ala Pro Ser Leu
        255                 260                 265

gag gac ggt tcg gat gcc ttc atg tca ccc cag gat gtt cgg ggc acc      928
Glu Asp Gly Ser Asp Ala Phe Met Ser Pro Gln Asp Val Arg Gly Thr
    270                 275                 280

tca gaa aac ctt cct gag agg tct gtc ccc tta cgc aag agc ctc tgc      976
Ser Glu Asn Leu Pro Glu Arg Ser Val Pro Leu Arg Lys Ser Leu Cys
285                 290                 295                 300

tcc ccc act ttc ctg tgg agc ctc ctc acc atg ggc atg acc cag ctg     1024
Ser Pro Thr Phe Leu Trp Ser Leu Leu Thr Met Gly Met Thr Gln Leu
                305                 310                 315

cgg atc atc ttc tac atg gct gct gtg aac aag atg ctg gag tac ctt     1072
Arg Ile Ile Phe Tyr Met Ala Ala Val Asn Lys Met Leu Glu Tyr Leu
            320                 325                 330

gtg act ggt ggc cag gag cat gag aca aat gaa cag caa caa aag gtg     1120
Val Thr Gly Gly Gln Glu His Glu Thr Asn Glu Gln Gln Gln Lys Val
```

```
          335                 340                 345

gca gag aca gtt ggg ttc tac tcc tcc gtc ttc ggg gcc atg cag ctg     1168
Ala Glu Thr Val Gly Phe Tyr Ser Ser Val Phe Gly Ala Met Gln Leu
    350                 355                 360

ttg tgc ctt ctc acc tgc ccc ctc att ggc tac atc atg gac tgg cgg     1216
Leu Cys Leu Leu Thr Cys Pro Leu Ile Gly Tyr Ile Met Asp Trp Arg
365                 370                 375                 380

atc aag gac tgc gtg gac gcc cca act cag ggc act gtc ctc gga gat     1264
Ile Lys Asp Cys Val Asp Ala Pro Thr Gln Gly Thr Val Leu Gly Asp
                385                 390                 395

gcc agg gac ggg gtt gct acc aaa tcc atc aga cca cgc tac tgc aag     1312
Ala Arg Asp Gly Val Ala Thr Lys Ser Ile Arg Pro Arg Tyr Cys Lys
            400                 405                 410

atc caa aag ctc acc aat gcc atc agt gcc ttc acc ctg acc aac ctg     1360
Ile Gln Lys Leu Thr Asn Ala Ile Ser Ala Phe Thr Leu Thr Asn Leu
        415                 420                 425

ctg ctt gtg ggt ttt ggc atc acc tgt ctc atc aac aac tta cac ctc     1408
Leu Leu Val Gly Phe Gly Ile Thr Cys Leu Ile Asn Asn Leu His Leu
    430                 435                 440

cag ttt gtg acc ttt gtc ctg cac acc att gtt cga ggt ttc ttc cac     1456
Gln Phe Val Thr Phe Val Leu His Thr Ile Val Arg Gly Phe Phe His
445                 450                 455                 460

tca gcc tgt ggg agt ctc tat gct gca gtg ttc cca tcc aac cac ttt     1504
Ser Ala Cys Gly Ser Leu Tyr Ala Ala Val Phe Pro Ser Asn His Phe
                465                 470                 475

ggg acg ctg aca ggc ctg cag tcc ctc atc agt gct gtg ttc gcc ttg     1552
Gly Thr Leu Thr Gly Leu Gln Ser Leu Ile Ser Ala Val Phe Ala Leu
            480                 485                 490

ctt cag cag cca ctt ttc atg gcg atg gtg gga ccc ctg aaa gga gag     1600
Leu Gln Gln Pro Leu Phe Met Ala Met Val Gly Pro Leu Lys Gly Glu
        495                 500                 505

ccc ttc tgg gtg aat ctg ggc ctc ctg cta ttc tca ctc ctg gga ttc     1648
Pro Phe Trp Val Asn Leu Gly Leu Leu Leu Phe Ser Leu Leu Gly Phe
    510                 515                 520

ctg ttg cct tcc tac ctc ttc tat tac cgt gcc cgg ctc cag cag gag     1696
Leu Leu Pro Ser Tyr Leu Phe Tyr Tyr Arg Ala Arg Leu Gln Gln Glu
525                 530                 535                 540

tac gcc gcc aat ggg atg ggc cca ctg aag gtg ctt agc ggc tct gag     1744
Tyr Ala Ala Asn Gly Met Gly Pro Leu Lys Val Leu Ser Gly Ser Glu
                545                 550                 555

gtg acc gca tagacttctc agaccaaggg acctggatga caggcaatca             1793
Val Thr Ala
```

```
aggcctgagc aaccaaaagg agtgccccat atggcttttc tacctgtaac atgcacatag   1853
agccatggcc gtagatttat aaataccaag agaagttcta tttttgtaaa gactgcaaaa   1913
aggaggaaaa aaaaccttca aaaacgcccc ctaagtcaac gctccattga ctgaagacag   1973
tccctatcct agaggggttg agctttcttc ctccttgggt tggaggagac cagggtgcct   2033
cttatctcct tctagcggtc tgcctcctgg tacctcttgg ggggatcggc aaacaggcta   2093
cccctgaggt cccatgtgcc atgagtgtgc acaacatgca atgtgtctgt gtatgtgtga   2153
atgtgagaaa aacacagccc tcctttcaga aggaaagggg cctgaggtgc cagctgtgtc   2213
ctgggttagg ggttgggggt cggccccttc cagggccagg aaggcaggtt ccctctctgg   2273
tgctgctgct tgcaagtctt agaggaaata aaaagggaag tgagaaaaaa aaa           2326
```

<210> 2
<211> 559
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Ala Pro Thr Leu Gln Gln Ala Tyr Arg Arg Arg Trp Trp Met Ala
 1               5                  10                  15
Cys Thr Ala Val Leu Glu Asn Leu Phe Phe Ser Ala Val Leu Leu Gly
            20                  25                  30
Trp Gly Ser Leu Leu Ile Ile Leu Lys Asn Glu Gly Phe Tyr Ser Ser
        35                  40                  45
Thr Cys Pro Ala Glu Ser Ser Thr Asn Thr Thr Gln Asp Glu Gln Arg
    50                  55                  60
Arg Trp Pro Gly Cys Asp Gln Gln Asp Glu Met Leu Asn Leu Gly Phe
65                  70                  75                  80
Thr Ile Gly Ser Phe Val Leu Ser Ala Thr Thr Leu Pro Leu Gly Ile
                85                  90                  95
Leu Met Asp Arg Phe Gly Pro Arg Pro Val Arg Leu Val Gly Ser Ala
            100                 105                 110
Cys Phe Thr Ala Ser Cys Thr Leu Met Ala Leu Ala Ser Arg Asp Val
        115                 120                 125
Glu Ala Leu Ser Pro Leu Ile Phe Leu Ala Leu Ser Leu Asn Gly Phe
    130                 135                 140
Gly Gly Ile Cys Leu Thr Phe Thr Ser Leu Thr Leu Pro Asn Met Phe
145                 150                 155                 160
Gly Asn Leu Arg Ser Thr Leu Met Ala Leu Met Ile Gly Ser Tyr Ala
                165                 170                 175
Ser Ser Ala Ile Thr Phe Pro Gly Ile Lys Leu Ile Tyr Asp Ala Gly
            180                 185                 190
Val Ala Phe Val Val Ile Met Phe Thr Trp Ser Gly Leu Ala Cys Leu
        195                 200                 205
Ile Phe Leu Asn Cys Thr Leu Asn Trp Pro Ile Glu Ala Phe Pro Ala
    210                 215                 220
Pro Glu Glu Val Asn Tyr Thr Lys Lys Ile Lys Leu Ser Gly Leu Ala
225                 230                 235                 240
Leu Asp His Lys Val Thr Gly Asp Leu Phe Tyr Thr His Val Thr Thr
                245                 250                 255
Met Gly Gln Arg Leu Ser Gln Lys Ala Pro Ser Leu Glu Asp Gly Ser
            260                 265                 270
Asp Ala Phe Met Ser Pro Gln Asp Val Arg Gly Thr Ser Glu Asn Leu
        275                 280                 285
Pro Glu Arg Ser Val Pro Leu Arg Lys Ser Leu Cys Ser Pro Thr Phe
    290                 295                 300
Leu Trp Ser Leu Leu Thr Met Gly Met Thr Gln Leu Arg Ile Ile Phe
305                 310                 315                 320
```

```
Tyr Met Ala Ala Val Asn Lys Met Leu Glu Tyr Leu Val Thr Gly Gly
                325                 330                 335
Gln Glu His Glu Thr Asn Glu Gln Gln Gln Lys Val Ala Glu Thr Val
            340                 345                 350
Gly Phe Tyr Ser Ser Val Phe Gly Ala Met Gln Leu Leu Cys Leu Leu
        355                 360                 365
Thr Cys Pro Leu Ile Gly Tyr Ile Met Asp Trp Arg Ile Lys Asp Cys
    370                 375                 380
Val Asp Ala Pro Thr Gln Gly Thr Val Leu Gly Asp Ala Arg Asp Gly
385                 390                 395                 400
Val Ala Thr Lys Ser Ile Arg Pro Arg Tyr Cys Lys Ile Gln Lys Leu
                405                 410                 415
Thr Asn Ala Ile Ser Ala Phe Thr Leu Thr Asn Leu Leu Leu Val Gly
            420                 425                 430
Phe Gly Ile Thr Cys Leu Ile Asn Asn Leu His Leu Gln Phe Val Thr
        435                 440                 445
Phe Val Leu His Thr Ile Val Arg Gly Phe Phe His Ser Ala Cys Gly
    450                 455                 460
Ser Leu Tyr Ala Ala Val Phe Pro Ser Asn His Phe Gly Thr Leu Thr
465                 470                 475                 480
Gly Leu Gln Ser Leu Ile Ser Ala Val Phe Ala Leu Leu Gln Gln Pro
                485                 490                 495
Leu Phe Met Ala Met Val Gly Pro Leu Lys Gly Glu Pro Phe Trp Val
            500                 505                 510
Asn Leu Gly Leu Leu Leu Phe Ser Leu Leu Gly Phe Leu Leu Pro Ser
        515                 520                 525
Tyr Leu Phe Tyr Tyr Arg Ala Arg Leu Gln Gln Glu Tyr Ala Ala Asn
    530                 535                 540
Gly Met Gly Pro Leu Lys Val Leu Ser Gly Ser Glu Val Thr Ala
545                 550                 555
```

<210> 3
<211> 3442
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1760)...(3439)

<400> 3

```
ccggggctgg agggggcaa gcgggttccg aggtgcaaag cctggtgccc cgagccctgc     60
ggagctcggg gccagcatgg cccccacgct gcaacaggcg taccggaggc gctggtggat    120
ggcctgcacg gctgtgctgg agaacctctt cttctctgct gtactcctgg gctggggctc    180
cctgttgatc attctgaaga acgagggctt ctattccagc acgtgcccag ctgagagcag    240
caccaacacc acccaggatg agcagcgcag gtggccaggc tgtgaccagc aggacgagat    300
gctcaacctg ggcttcacca ttggttcctt cgtgctcagc gccaccaccc tgccactggg    360
gatcctcatg gaccgctttg gcccccgacc cgtgcggctg gttggcagtg cctgcttcac    420
tgcgtcctgc accctcatgg ccctggcctc ccgggacgtg gaagctctgt ctccgttgat    480
attcctggcg ctgtccctga atggctttgg tggcatctgc ctaacgttca cttcactcac    540
gctgcccaac atgtttggga acctgcgctc cacgttaatg gccctcatga ttggctctta    600
cgcctcttct gccattacgt tcccaggaat caagctgatc tacgatgccg gtgtggcctt    660
cgtggtcatc atgttcacct ggtctggcct ggcctgcctt atctttctga actgcaccct    720
caactggccc atcgaagcct ttcctgcccc tgaggaagtc aattacacga agaagatcaa    780
gctgagtggg ctggccctgg accacaaggt gacaggtgac ctcttctaca cccatgtgac    840
caccatgggc cagaggctca gccagaaggc ccccagcctg gaggacggtt cggatgcctt    900
catgtcaccc caggatgttc ggggcacctc agaaaacctt cctgagaggt ctgtcccctt    960
```

```
acgcaagagc ctctgctccc ccactttcct gtggagcctc ctcaccatgg gcatgaccca   1020
gctgcggatc atcttctaca tggctgctgt gaacaagatg ctggagtacc ttgtgactgg   1080
tggccaggag catgagacaa atgaacagca acaaaaggtg gcagagacag ttgggttcta   1140
ctcctccgtc ttcggggcca tgcagctgtt gtgccttctc acctgccccc tcattggcta   1200
catcatggac tggcggatca aggactgcgt ggacgcccca actcagggca ctgtcctcgg   1260
agatgccagg gacggggttg ctaccaaatc catcagacca cgctactgca agatccaaaa   1320
gctcaccaat gccatcagtg ccttcaccct gaccaacctg ctgcttgtgg gttttggcat   1380
cacctgtctc atcaacaact tacacctcca gtttgtgacc tttgtcctgc acaccattgt   1440
tcgaggtttc ttccactcag cctgtgggag tctctatgct gcagtgttcc catccaacca   1500
ctttgggacg ctgacaggcc tgcagtccct catcagtgct gtgttcgcct tgcttcagca   1560
gccacttttc atggcgatgg tgggaccccc gaaaggagag cccttctggg tgagagcgag   1620
ggttggtgtg gggggagcag gagccactct cctgggggca ggggtagggc cttgtatgtg   1680
gtgccatccc tcactcatct cagccagagg cacctcagag gtctctaatc tgcaggtttc   1740
caagttgtct gccttttag atg gcc ccc acg ctg caa cag gcg tac cgg agg     1792
                    Met Ala Pro Thr Leu Gln Gln Ala Tyr Arg Arg
                     1               5                  10

cgc tgg tgg atg gcc tgc acg gct gtg ctg gag aac ctc ttc ttc tct     1840
Arg Trp Trp Met Ala Cys Thr Ala Val Leu Glu Asn Leu Phe Phe Ser
            15                  20                  25

gct gta ctc ctg ggc tgg ggc tcc ctg ttg atc att ctg aag aac gag     1888
Ala Val Leu Leu Gly Trp Gly Ser Leu Leu Ile Ile Leu Lys Asn Glu
        30                  35                  40

ggc ttc tat tcc agc acg tgc cca gct gag agc agc acc aac acc acc     1936
Gly Phe Tyr Ser Ser Thr Cys Pro Ala Glu Ser Ser Thr Asn Thr Thr
    45                  50                  55

cag gat gag cag cgc agg tgg cca ggc tgt gac cag cag gac gag atg     1984
Gln Asp Glu Gln Arg Arg Trp Pro Gly Cys Asp Gln Gln Asp Glu Met
60                  65                  70                  75

ctc aac ctg ggc ttc acc att ggt tcc ttc gtg ctc agc gcc acc acc     2032
Leu Asn Leu Gly Phe Thr Ile Gly Ser Phe Val Leu Ser Ala Thr Thr
                80                  85                  90

ctg cca ctg ggg atc ctc atg gac cgc ttt ggc ccc cga ccc gtg cgg     2080
Leu Pro Leu Gly Ile Leu Met Asp Arg Phe Gly Pro Arg Pro Val Arg
            95                  100                 105

ctg gtt ggc agt gcc tgc ttc act gcg tcc tgc acc ctc atg gcc ctg     2128
Leu Val Gly Ser Ala Cys Phe Thr Ala Ser Cys Thr Leu Met Ala Leu
        110                 115                 120

gcc tcc cgg gac gtg gaa gct ctg tct ccg ttg ata ttc ctg gcg ctg     2176
Ala Ser Arg Asp Val Glu Ala Leu Ser Pro Leu Ile Phe Leu Ala Leu
    125                 130                 135

tcc ctg aat ggc ttt ggt ggc atc tgc cta acg ttc act tca ctc acg     2224
Ser Leu Asn Gly Phe Gly Gly Ile Cys Leu Thr Phe Thr Ser Leu Thr
140                 145                 150                 155

ctg ccc aac atg ttt ggg aac ctg cgc tcc acg tta atg gcc ctc atg     2272
Leu Pro Asn Met Phe Gly Asn Leu Arg Ser Thr Leu Met Ala Leu Met
                160                 165                 170
```

```
att ggc tct tac gcc tct tct gcc att acg ttc cca gga atc aag ctg      2320
Ile Gly Ser Tyr Ala Ser Ser Ala Ile Thr Phe Pro Gly Ile Lys Leu
            175                 180                 185

atc tac gat gcc ggt gtg gcc ttc gtg gtc atc atg ttc acc tgg tct      2368
Ile Tyr Asp Ala Gly Val Ala Phe Val Val Ile Met Phe Thr Trp Ser
        190                 195                 200

ggc ctg gcc tgc ctt atc ttt ctg aac tgc acc ctc aac tgg ccc atc      2416
Gly Leu Ala Cys Leu Ile Phe Leu Asn Cys Thr Leu Asn Trp Pro Ile
    205                 210                 215

gaa gcc ttt cct gcc cct gag gaa gtc aat tac acg aag aag atc aag      2464
Glu Ala Phe Pro Ala Pro Glu Glu Val Asn Tyr Thr Lys Lys Ile Lys
220                 225                 230                 235

ctg agt ggg ctg gcc ctg gac cac aag gtg aca ggt gac ctc ttc tac      2512
Leu Ser Gly Leu Ala Leu Asp His Lys Val Thr Gly Asp Leu Phe Tyr
                240                 245                 250

acc cat gtg acc acc atg ggc cag agg ctc agc cag aag gcc ccc agc      2560
Thr His Val Thr Thr Met Gly Gln Arg Leu Ser Gln Lys Ala Pro Ser
            255                 260                 265

ctg gag gac ggt tcg gat gcc ttc atg tca ccc cag gat gtt cgg ggc      2608
Leu Glu Asp Gly Ser Asp Ala Phe Met Ser Pro Gln Asp Val Arg Gly
        270                 275                 280

acc tca gaa aac ctt cct gag agg tct gtc ccc tta cgc aag agc ctc      2656
Thr Ser Glu Asn Leu Pro Glu Arg Ser Val Pro Leu Arg Lys Ser Leu
    285                 290                 295

tgc tcc ccc act ttc ctg tgg agc ctc ctc acc atg ggc atg acc cag      2704
Cys Ser Pro Thr Phe Leu Trp Ser Leu Leu Thr Met Gly Met Thr Gln
300                 305                 310                 315

ctg cgg atc atc ttc tac atg gct gct gtg aac aag atg ctg gag tac      2752
Leu Arg Ile Ile Phe Tyr Met Ala Ala Val Asn Lys Met Leu Glu Tyr
                320                 325                 330

ctt gtg act ggt ggc cag gag cat gag aca aat gaa cag caa caa aag      2800
Leu Val Thr Gly Gly Gln Glu His Glu Thr Asn Glu Gln Gln Gln Lys
            335                 340                 345

gtg gca gag aca gtt ggg ttc tac tcc tcc gtc ttc ggg gcc atg cag      2848
Val Ala Glu Thr Val Gly Phe Tyr Ser Ser Val Phe Gly Ala Met Gln
        350                 355                 360

ctg ttg tgc ctt ctc acc tgc ccc ctc att ggc tac atc atg gac tgg      2896
Leu Leu Cys Leu Leu Thr Cys Pro Leu Ile Gly Tyr Ile Met Asp Trp
    365                 370                 375

cgg atc aag gac tgc gtg gac gcc cca act cag ggc act gtc ctc gga      2944
Arg Ile Lys Asp Cys Val Asp Ala Pro Thr Gln Gly Thr Val Leu Gly
380                 385                 390                 395

gat gcc agg gac ggg gtt gct acc aaa tcc atc aga cca cgc tac tgc      2992
```

```
Asp Ala Arg Asp Gly Val Ala Thr Lys Ser Ile Arg Pro Arg Tyr Cys
                400                 405                 410

aag atc caa aag ctc acc aat gcc atc agt gcc ttc acc ctg acc aac     3040
Lys Ile Gln Lys Leu Thr Asn Ala Ile Ser Ala Phe Thr Leu Thr Asn
            415                 420                 425

ctg ctg ctt gtg ggt ttt ggc atc acc tgt ctc atc aac aac tta cac     3088
Leu Leu Leu Val Gly Phe Gly Ile Thr Cys Leu Ile Asn Asn Leu His
        430                 435                 440

ctc cag ttt gtg acc ttt gtc ctg cac acc att gtt cga ggt ttc ttc     3136
Leu Gln Phe Val Thr Phe Val Leu His Thr Ile Val Arg Gly Phe Phe
    445                 450                 455

cac tca gcc tgt ggg agt ctc tat gct gca gtg ttc cca tcc aac cac     3184
His Ser Ala Cys Gly Ser Leu Tyr Ala Ala Val Phe Pro Ser Asn His
460                 465                 470                 475

ttt ggg acg ctg aca ggc ctg cag tcc ctc atc agt gct gtg ttc gcc     3232
Phe Gly Thr Leu Thr Gly Leu Gln Ser Leu Ile Ser Ala Val Phe Ala
                480                 485                 490

ttg ctt cag cag cca ctt ttc atg gcg atg gtg gga ccc ctg aaa gga     3280
Leu Leu Gln Gln Pro Leu Phe Met Ala Met Val Gly Pro Leu Lys Gly
            495                 500                 505

gag ccc ttc tgg gtg aga gcg agg gtt ggt gtg ggg gga gca gga gcc     3328
Glu Pro Phe Trp Val Arg Ala Arg Val Gly Val Gly Gly Ala Gly Ala
        510                 515                 520

act ctc ctg ggg gca ggg gta ggg cct tgt atg tgg tgc cat ccc tca     3376
Thr Leu Leu Gly Ala Gly Val Gly Pro Cys Met Trp Cys His Pro Ser
    525                 530                 535

ctc atc tca gcc aga ggc acc tca gag gtc tct aat ctg cag gtt tcc     3424
Leu Ile Ser Ala Arg Gly Thr Ser Glu Val Ser Asn Leu Gln Val Ser
540                 545                 550                 555

aag ttg tct gcc ttt tag                                             3442
Lys Leu Ser Ala Phe
                560
```

<210> 4
<211> 560
<212> PRT
<213> Homo sapiens

<400> 4

```
Met Ala Pro Thr Leu Gln Gln Ala Tyr Arg Arg Arg Trp Trp Met Ala
 1               5                  10                  15
Cys Thr Ala Val Leu Glu Asn Leu Phe Phe Ser Ala Val Leu Leu Gly
            20                  25                  30
Trp Gly Ser Leu Leu Ile Ile Leu Lys Asn Glu Gly Phe Tyr Ser Ser
        35                  40                  45
Thr Cys Pro Ala Glu Ser Ser Thr Asn Thr Thr Gln Asp Glu Gln Arg
```

```
    50                  55                  60
Arg Trp Pro Gly Cys Asp Gln Gln Asp Glu Met Leu Asn Leu Gly Phe
65                  70                  75                  80
Thr Ile Gly Ser Phe Val Leu Ser Ala Thr Thr Leu Pro Leu Gly Ile
                85                  90                  95
Leu Met Asp Arg Phe Gly Pro Arg Pro Val Arg Leu Val Gly Ser Ala
            100                 105                 110
Cys Phe Thr Ala Ser Cys Thr Leu Met Ala Leu Ala Ser Arg Asp Val
        115                 120                 125
Glu Ala Leu Ser Pro Leu Ile Phe Leu Ala Leu Ser Leu Asn Gly Phe
    130                 135                 140
Gly Gly Ile Cys Leu Thr Phe Thr Ser Leu Thr Leu Pro Asn Met Phe
145                 150                 155                 160
Gly Asn Leu Arg Ser Thr Leu Met Ala Leu Met Ile Gly Ser Tyr Ala
                165                 170                 175
Ser Ser Ala Ile Thr Phe Pro Gly Ile Lys Leu Ile Tyr Asp Ala Gly
            180                 185                 190
Val Ala Phe Val Val Ile Met Phe Thr Trp Ser Gly Leu Ala Cys Leu
        195                 200                 205
Ile Phe Leu Asn Cys Thr Leu Asn Trp Pro Ile Glu Ala Phe Pro Ala
    210                 215                 220
Pro Glu Glu Val Asn Tyr Thr Lys Lys Ile Lys Leu Ser Gly Leu Ala
225                 230                 235                 240
Leu Asp His Lys Val Thr Gly Asp Leu Phe Tyr Thr His Val Thr Thr
                245                 250                 255
Met Gly Gln Arg Leu Ser Gln Lys Ala Pro Ser Leu Glu Asp Gly Ser
            260                 265                 270
Asp Ala Phe Met Ser Pro Gln Asp Val Arg Gly Thr Ser Glu Asn Leu
        275                 280                 285
Pro Glu Arg Ser Val Pro Leu Arg Lys Ser Leu Cys Ser Pro Thr Phe
    290                 295                 300
Leu Trp Ser Leu Leu Thr Met Gly Met Thr Gln Leu Arg Ile Ile Phe
305                 310                 315                 320
Tyr Met Ala Ala Val Asn Lys Met Leu Glu Tyr Leu Val Thr Gly Gly
                325                 330                 335
Gln Glu His Glu Thr Asn Glu Gln Gln Gln Lys Val Ala Glu Thr Val
            340                 345                 350
Gly Phe Tyr Ser Ser Val Phe Gly Ala Met Gln Leu Leu Cys Leu Leu
        355                 360                 365
Thr Cys Pro Leu Ile Gly Tyr Ile Met Asp Trp Arg Ile Lys Asp Cys
    370                 375                 380
Val Asp Ala Pro Thr Gln Gly Thr Val Leu Gly Asp Ala Arg Asp Gly
385                 390                 395                 400
Val Ala Thr Lys Ser Ile Arg Pro Arg Tyr Cys Lys Ile Gln Lys Leu
                405                 410                 415
Thr Asn Ala Ile Ser Ala Phe Thr Leu Thr Asn Leu Leu Leu Val Gly
            420                 425                 430
Phe Gly Ile Thr Cys Leu Ile Asn Asn Leu His Leu Gln Phe Val Thr
        435                 440                 445
Phe Val Leu His Thr Ile Val Arg Gly Phe Phe His Ser Ala Cys Gly
    450                 455                 460
Ser Leu Tyr Ala Ala Val Phe Pro Ser Asn His Phe Gly Thr Leu Thr
465                 470                 475                 480
Gly Leu Gln Ser Leu Ile Ser Ala Val Phe Ala Leu Leu Gln Gln Pro
                485                 490                 495
Leu Phe Met Ala Met Val Gly Pro Leu Lys Gly Glu Pro Phe Trp Val
            500                 505                 510
```

Arg Ala Arg Val Gly Val Gly Gly Ala Gly Ala Thr Leu Leu Gly Ala
515 520 525
Gly Val Gly Pro Cys Met Trp Cys His Pro Ser Leu Ile Ser Ala Arg
530 535 540
Gly Thr Ser Glu Val Ser Asn Leu Gln Val Ser Lys Leu Ser Ala Phe
545 550 555 560

<210> 5
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Arbitrary primer A2 from Stratagene, Inc.

<400> 5
aatctagagc tccagcag 18

<210> 6
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Zinc finger-directed primer

<400> 6
gtcgtcgaat tccacacagg agaaaagcc 29

<210> 7
<211> 22
<212> DNA
<213> Homo sapiens

<400> 7
gcatgttaca ggtagaaaag cc 22

<210> 8
<211> 21
<212> DNA
<213> Homo sapiens

<400> 8
ctggcgtatc tgaagagtct g 21

<210> 9
<211> 103
<212> DNA
<213> Homo sapiens

<400> 9
acaggaatcc ccaggagtga agaataagca ggaggcccca gattcacctt tagggcaagg 60
agagagaaac agagtcaagt aggtagtcat ctgcccttaa gcc 103

<210> 10
<211> 20
<212> DNA

<213> Homo sapiens

<400> 10
gaccgcatag acttctcaga 20

<210> 11
<211> 22
<212> DNA
<213> Homo sapiens

<400> 11
tctgcaaagt ggctgagatg ag 22

<210> 12
<211> 26
<212> DNA
<213> Homo sapiens

<400> 12
cctgccttat ctttctgaac tgcacc 26

<210> 13
<211> 14
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope

<400> 13
Thr Gln Asp Glu Gln Arg Arg Trp Pro Gly Cys Asp Gln Gln
1 5 10

<210> 14
<211> 14
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope

<400> 14
Glu Asn Leu Pro Glu Arg Ser Val Pro Leu Arg Lys Ser Leu
1 5 10

<210> 15
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope

<400> 15
Arg Pro Arg Tyr Cys Lys Ile Gln Lys Leu Thr Asn Ala
1 5 10

<210> 16
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope

<400> 16
Ala Asn Gly Met Gly Pro Leu Lys Val Leu Ser Gly Ser
1 5 10

<210> 17
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope

<400> 17
Ala Arg Gly Thr Ser Glu Val Ser Asn Leu Gln Val Ser
1 5 10

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/16831

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C07K16/18 C12Q1/68 G01N33/53

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C12Q C07K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	COLE ET AL.: "cDNA sequencing and analysis of POV1 (PB39): a novel gene up-regulated in prostate cancer" GENOMICS, vol. 51, no. 2, July 1998 (1998-07), pages 282-287, XP002120985 page 282, column 2 -page 287, column 1; figures 1,2 ---	1-16
X	CHUAQUI ET AL.: "Identification of a novel transcript up-regulated in a clinically aggressive prostate carcinoma" UROLOGY, vol. 50, no. 2, August 1997 (1997-08), pages 302-307, XP002120986 page 304 -page 306; figures 1,3; table 1 ---	1-16

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 November 1999	17/11/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	van Klompenburg, W

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/16831

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HILLIER ET AL.: "The WashU-Merck EST Project" EMBL ACC NO: R00504, 17 April 1995 (1995-04-17), XP002120987 the whole document	1-8
X	WO 98 21328 A (KATO SEISHI ;PROTEGENE INC (JP); SEKINE SHINGO (JP); SAGAMI CHEM R) 22 May 1998 (1998-05-22) SEQ ID NO 71 page 166 -page 169; claim 4	1-8
A	WO 98 10098 A (FISHER PAUL B ;UNIV COLUMBIA (US)) 12 March 1998 (1998-03-12) page 10, line 17 - line 25; claims 1-20; figures 1,4	1-16

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 99/16831

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9821328 A	22-05-1998	AU 4885297 A	03-06-1998
		EP 0941320 A	15-09-1999
WO 9810098 A	12-03-1998	AU 4182697 A	26-03-1998
		EP 0951567 A	27-10-1999

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☑ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.